THE STRUCTURE AND INFRASTRUCTURE OF THE FINNISH RESEARCH LITERATURE

By

Dr. Ronald N. Kostoff, Office of Naval Research, 875 N. Randolph St., Arlington, VA 22217
Phone: 703-696-4198; Fax: 703-696-8744; Internet: kostofr@onr.navy.mil

Mr. Rene Tshiteya, DDL-OMNI Engineering, LLC, 8260 Greensboro Drive, Suite 600, Mclean, VA 22102, Presently at Koniag Services Inc

Ms. Christine A. Bowles, DDL-OMNI Engineering, LLC,
8260 Greensboro Drive, Suite 600, Mclean, VA 22102

Dr. Tuure Tuunanen, University of Auckland Business School,
The Department of ISOM, Private Bag 92019, Auckland, New Zealand

(*The views in this report are solely those of the authors, and do not necessarily represent the views of the U.S. Department of the Navy or any of its components, DDL-OMNI, Koniag Services, or University of Auckland Business School, New Zealand*)

KEYWORDS

Finland; Science and Technology; Bibliometrics; Citation Analysis; Computational Linguistics; Core Competencies; Research Evaluation; Factor Analysis; Concept Clustering; Document Clustering; CLUTO; Text Mining; Wireless Networks; Mobile Communications; Signal Processing.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
06 FEB 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
The Structure and Infrastructure of the Finnish Research Literature	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Office of Naval Research Dr. Ronald Kostoff 875 N. Randolph Street Arlington, VA 22217**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**SAR**	**98**	

ABSTRACT

The structure and infrastructure of the Finnish research literature was determined. A representative database of technical articles was extracted from the Science Citation Index for the years 2003-2004, with each article containing at least one author with a Finnish address. Document clustering was used to identify the main technical themes (core competencies) of Finnish research. Four of the pervasive technical topics identified from the clustering (Wireless Networks and Mobile Communication, Signal Processing, Materials Science and Engineering, Chemistry) were analyzed further using bibliometrics, in order to identify the infrastructure of these research areas. Finally, the citation performance of Finnish research in the four pervasive technical topics above, and in other technical topics obtained by analysis of Abstract phrases, was compared to that of two Scandanavian countries with similar population and GDP: Norway and Denmark.

BACKGROUND

The Background section addresses four major components of this study.

- Core competencies, from their business origins to their national research context, since Finland's core competencies are one of the focal points of the present assessment.
- Country technology assessments, including both strengths and weaknesses of the major approaches.
- Text mining and its potential role in country technology assessments.
- Finland's science and technology structure, to provide a context into which the results and discussion can be placed.

Core Competencies

The core competence concept was initially promulgated in 1990 as "an area of specialized expertise that is the result of harmonizing complex streams of technology and work activity" (Hamel and Prahalad, 1990). It was developed for a business context, and reflected the collective learning and coordination skills underlying a firm's product lines. According to the original proposers, core competencies are the source of competitive advantage and enable the firm to introduce an array of new products and services. They lead to the development of core products, which are then used to develop a larger number of end user products.

Since the original core competence article, many follow-on studies have been performed. Other definitions of core competence have been advanced (e.g., Galunic and Rodan, 1998). However, common features among the different core competence definitions include the following:

- Critical mass of people
- Synergy of coordinated sub-disciplines
- High quality output
- Unique capabilities
- Substantial fraction of organization's total development investment

While the original definition and most follow-on definitions have applied to business organizations, the concept can be extrapolated to nations. The five

features above characterize national core competencies. In the present report, a national research core competence is defined as a technical area that

- contains a critical mass of researchers;
- consists of coordinated and synchronized sub-disciplines;
- produces high quality output;
- offers unique national capabilities;
- contains a visible fraction of research investment.

In other words, a national research core competence is a synergy of individual expertise that is aggregated and coordinated over multiple technical disciplines and is expressed as a national research strategic investment.

The text mining approach of the present report address a sub-set of the above features (identification of Finland's main research thrusts, volume of research output in main research thrusts, relative quality of selected major research thrusts) to assess potential Finnish research competencies. Further subjective analysis (beyond the scope of the present report) is required to characterize the remaining necessary features of a national core competence.

This report will not discuss the desirability of employing core competencies in managing research. The first author has consulted with companies and agencies on practical aspects of implementing core competencies in research management. Within an organization, development of research core competencies tends to receive preferential and protected funding, which are very important in times of economic turndown. *Serious employee morale problems can result for those researchers who are not associated with core competence development, since they have been placed in a more vulnerable position.* The alternative, defining all the organization's development thrusts as core competencies, dilutes the purpose of utilizing core competencies to help manage research and renders them ineffective.

Country Technology Assessments

National science and technology (S&T) core competencies represent a country's strategic capabilities in S&T. Knowledge of country core competencies is important for myriad reasons:

- Identification of Priority technical areas for joint commercial or military ventures
- Assessment of a country's military potential
- Knowledge of emerging areas to avoid commercial or military surprise

Formatted: Bullets and Numbering

Obtaining such global technical awareness, especially from the literature, is difficult for multiple reasons:

- Much science and technology performed is not documented
- Much documented science and technology is not widely available
- Much available documented science and technology is expensive and difficult to acquire
- Few credible techniques exist for extracting useful information from large amounts of science and technology documentation (Kostoff, 2003)

Formatted: Bullets and Numbering

Most credible country technology assessments have been based on a combination of personal visitations to the country of interest, supplemented by copious reading of technology reports from that country. Such processes tend to be laborious, slow, expensive, and accompanied by large gaps in the knowledge available. The more credible and complete evaluation processes have focused on selected technologies from a particular country and provided in-depth analysis.

For the past half century, driven mainly by the Cold War, a large number of country technology assessments were performed (Bostian et al (2000), Leneman (1984), Stares (1985), Hutubessy et al (2002), Mooney and Seymour (1996), McIntire (2003), Campbell et al (1985), Klinger (1990), Cohn et al (1993), Lanzerotti et al (1986), Duncan et al (1988), Spencer et al (1989), Davidson et al (1990)). The last decade has seen an expansion in focus to technologies of major economic competitors. Over the past two decades, some of the most credible of these country technology assessments have come from two organizations: World Technology Evaluation Center (WTEC-Loyola Univ) and Foreign Applied Sciences Assessment Center (FASAC-SAIC). In conducting their studies, both of these organizations would gather topical literature from the country of interest, assemble teams of experts in the topical area, have the teams review the literature as well as conduct site visitations, and have the teams brief their findings and write a

final report. The studies performed by these groups remain seminal approaches to country technology assessments.

Text Mining Technology Assessments

The first author's group has been developing text mining approaches to extract useful information from the global science and technology literature for the past decade (Kostoff (1997), Kostoff and DeMarco (2001), Kostoff et al (1998a, 1999, 2000a, 2000b, 2001, 2002, 2004a, 2004b, 2004c, 2005a)). These studies have typically focused on a technical discipline, and have examined global S&T efforts in this discipline. It is believed that such approaches, with slight modification, could be adapted to identifying the core S&T competencies in selected countries or regions, including estimation of the relative levels of effort in each of the core technology areas. It is also believed that coupling of the text mining approach with WTEC and FASAC approaches would amplify the strengths of each approach and reduce the limitations. The text mining component would be performed initially to identify:

- Key core competencies and technology thrusts in the country of interest
- Key interdisciplinary thrusts
- Approximate levels of effort in technology-specific competency areas and in interdisciplinary areas
- Highly productive researchers
- Highly productive Centers of Excellence, including those not well known
- Highly cited researchers

Once the key technologies, researchers, and Centers of Excellence had been identified, then site visitation strategies could be developed. The second phase of the effort would be the actual site visitations. A key step in this hybrid process would be demonstration of the ability of text mining to identify the targets of interest with reasonable precision in a timely manner at an acceptable cost. These three driving parameters (performance, time, cost) could be traded-off against each other to provide a balance acceptable and tailored to a variety of potential customers.

An initial country application of the first author's text mining approach assessed the core research competencies of Mexico (Kostoff et al, 2005b). This study examined myriad clustering techniques for identifying core

competencies. As a precursor to visits by ONR personnel to Finland, it was desired to identify core Finland research competencies, using the experience gained from the Mexico study, and developing additional capabilities. This would allow an efficient site visitation strategy to be generated.

The present study incorporates the site visitation information into a larger portrayal of diverse facets of Finnish research. After a summary overview of Finland's S&T structure, the bibliometrics and computational linguistics of Finland's recent research output are shown. These computations provide the technical infrastructure and structure of Finland's research. Then, with the use of citation analysis, the pervasive research thrusts of Finland are compared with those of Norway and Denmark to assess relative level of effort and impact of research. Only one clustering technique was selected, based on the Mexico clustering results.

Finnish Science and Technology Structure

Finland is a nation of 5 million people and is one of the Nordic countries residing in the northern part of Europe. The government has emphasized the importance of research and science during the last two decades. Finland increased its research and science funding by 52% during years 1977-2000 (Rantanen, 2004). Nokia, as a global player in the telecom markets, has been one of the most visible results of this agenda. Since year 2000, the S & T expenditures have been 3.4-3.5% of the Gross Domestic Product (GDP).

Finnish research funding is channeled through several organizations under governmental control. Universities receive approximately 30% of direct funding through their annual budget. Finnish Academia and National Technology Agency of Finland (Tekes) are governmental research agencies that provide external funding to research projects. Tekes supports applied research programs, such as mobile technology and biotechnology, whereas Finnish Academy's role is to provide more strategic funding for basic research. In year 2005, these institutes provided about 40% of the total S & T funding in Finland. In addition, there are governmental research institutes that receive direct funding from the government. These institutes, like VTT research, received about 14% of total S & T funding in year 2005. The emphasis of the governmental research support has recently shifted to provide spearhead funding for creating international level research units. Finnish Academy is a key player in this process. Since year 2000, it has been running a Centre of Excellence programme. The strategic aim of the

programme is to provide additional funding to help the internationalization efforts of top Finnish research units. Currently, the programme consists of 26 research units divided among Finnish research universities (Finnish Academy, 2005).

Finnish university structure is divided into 11 research universities and 31 polytechnical universities (Rantanen, 2004). However, the research universities also have about 60 local establishments spread throughout the country. Hence, many of the medium size cities (30-50 000 residents) have a university representation in some form. There were about 21 000 academic researchers working in the research universities in year 2004 and some 11 000 teachers or other staff in the polytechnical universities. Currently, the polytechnical schools are recruiting additional Ph.Ds to their staff. Hence, the research will likely get more emphasis in this sector also. Rantanen (2004) has estimated that with this university system and with Finnish governmental research institutes, Finnish researchers contributed about 1% of the global research input in year 2002. Additionally, Rantanen estimated that Finnish researchers received 1.15% of all citations in the world during the same year. In addition to the university sector, there are about 10 000 researchers employed by the public sector and around 40 000 researchers working in commercial sector (Statistics Finland, 2004).

The Ministry of Education (The Ministery of Education, 2003) has given strategic emphasis to electrical and electronic engineering and biotechnology research for the on-going planning period of 2003-2006. Furthermore, the ministery has outlined the strategic importance of supporting commercialization of research findings and creation of new business areas through this process. The strategic objective is to secure the continuity of the positive development of Finnish academic research.

Objectives

Identify the S&T core competencies of Finland. Compare a few competencies with those of other countries that have been studied. Further, generate a process that could be used efficiently and rapidly to assess the S&T core competencies in other countries of interest.

APPROACH AND RESULTS

1. Overview

In a study based on analysis of the research literature, two major types of information are required for a country S&T core competency assessment. One is technical infrastructure, which encompasses the prolific authors, journals that contain many of the papers, the prolific institutions, and the most cited papers/ authors/ journals. The other is technology thrusts and the relationship among the thrusts. This study focused on obtaining both types of information.

Two types of results are presented: bibliometrics and taxonomies. Bibliometrics provide an indication of the technical infrastructure (prolific authors, journals, institutions, citations), while taxonomies provide an indication of major technology thrusts and their relationships.

Section 2 describes the database used for the bibliometrics and taxonomy analyses. Section 3 presents the approaches and results, where

- section 3.1 presents the publication bibliometrics,
- section 3.2 presents the citation bibliometrics, and
- section 3.3 presents the document clustering taxonomy approach and results.

2. Databases and Information Retrieval Approach

For the present study, the Science Citation Index (SCI) database was used. The SCI accesses the premier basic and applied research journals, and analysis of the retrieved data will help identify the research core competencies of Finland. The retrieved database used for analysis consists of selected journal records (including the fields of authors, titles, journals, author addresses, author keywords, abstract narratives, and references cited for each paper) obtained by searching the Web version of the SCI for articles that contained at least one author with a Finland address. At the time the final data was extracted for the present report (Fall 2004), the expanded version of the SCI accessed about 5600 journals (mainly in physical, engineering, environmental, and life sciences basic research).

3. Bibliometrics

Retrospectively, there were 15767 research articles with at least one Finnish author address published in total 2003-2004. At the time of retrieval (Fall 2004), the then 12900 records that had already been published in 2003-2004 were retrieved, and the bibliometrics data were extracted.

3.1 Publication Statistics on Authors, Journals, and Organizations

Overview

The first group of metrics presents counts of papers published by different entities. These metrics can be viewed as output and productivity measures. They are not direct measures of research quality, although there is some threshold quality level inferred since these papers are published in the (typically) high caliber journals accessed by the SCI.

In all previous text mining studies published by the first author's group, bibliometrics were performed on the overall database retrieved. Since all these previous studies focused on a specific technology (with the exception of Kostoff et al, 2005b), the resultant bibliometrics provided the technical infrastructure for that technology. In the present case, the focus is on the wide range of technologies being developed within a country. Applying the bibliometrics analysis to the total retrieved database for that country will not provide very useful results. Visitation strategies (one desired application) are typically developed for a specific technology using a group of experts for that technology.

The approach taken here is to identify the thematic thrust areas from the document clustering performed in the latter part of this report, then retrieve documents that address each theme. The bibliometrics will then be performed on a theme by theme basis. For the present study, one theme is selected as an illustrative example for the bibliometrics in the main body of the text, and three other themes' bibliometrics are presented in Appendix 1.

Many thrust areas, at different levels of aggregation, can be identified from the clustering schematic in the latter part of this report. For illustration, the authors selected four thrust areas as potential competencies, based on the cohesiveness of the theme and the volume of the publications contained within the thrust. Two relatively specific areas were selected (Wireless Networks and Mobile Communications; Signal Processing), and two more aggregate level areas were also selected (Materials Science and Engineering;

Chemistry). Only Physical and Engineering Sciences topics were selected; Environmental Research or Life Sciences thrusts were not selected from the document clustering portion of the present analysis. Environmental research, such as forestry or species research, did tend to be very country-specific, and would appear to fulfill one of the characteristics of core competence: uniqueness. However, it was decided to focus on research areas in which Finland had a presence and which had more of an international scope. The same reasoning was used for many of the medical research topics, which focused on diseases prevalent in Finland. An expanded study would include some of the more basic biomedical topics, such as genetics.

In addition, a citation-based approach was used to identify pervasive research thrusts in Finland, and compare Finnish investment and impact with those of other countries. This approach is described in detail in section 3.2, citations. Basically, this approach identifies high frequency technical phrases from analysis of the retrieved Finland records, retrieves SCI records using selected phrases, and examines citation metrics from these records relative to those from similar countries. Both Environmental and Life Sciences records/ themes were included in this analysis.

Publication Bibliometrics for WNMC

Based on the computational linguistics (document clustering) results, Wireless Networks and Mobile Communications (WNMC) is an important thrust area of Finnish research and has a critical mass of effort. Starting with the phrases generated by the clustering algorithm for the WNMC cluster, an iterative feedback approach was used to generate the following comprehensive query for this research in Finland:

WIRELESS OR WLAN OR WPAN OR WCDMA OR PACKET RADIO NETWORK OR BLUETOOTH OR ALOHA SYSTEM OR CDMA OR TDMA OR DS-CDMA OR CHANNEL LOAD SENSING PROTOCOL OR CLSP OR MOBILE DEVICE* OR (MOBILE SAME (ACCESS CONTROL OR MULTIPLE ACCESS OR CODE DIVISION OR CHANNEL ESTIMATION OR FADING OR NETWORK* OR DEVICE* OR COMMUNICAT* OR PHONE* OR TERMINAL* OR INTERNET)) OR ((COMMUNICATION* OR NETWORK*) SAME (AD HOC OR GSM OR IR OR INFRARED OR RADIO OR IEEE 802* OR BANDWIDTH OR CHANNEL* OR RECEIVER* OR MAC OR MEDIUM ACCESS CONTROL OR UWB))

The query was inserted into the Science Citation Index, and about 210 records were recovered for the period 2003-2004. The bibliometrics analysis example was performed on these records.

3.1.1. Prolific Authors

Table 1 lists the fourteen most prolific authors in Finnish WNMC research, including their institutions. These were followed by fifteen authors that had three publications (not shown). Four institutions predominate, in the following order: Helsinki University of Technology, Tampere University of Technology, Oulu University, University of Turku.

TABLE 1 – MOST PROLIFIC FINNISH WNMC AUTHORS

AUTHOR	INSTITUTION	#PAPERS
VAINIKAINEN--P	HELSINKI UNIV TECHNOLOGY	15
LATVA-AHO--M	OULU UNIV	8
HANNIKAINEN--M	TAMPERE UNIV TECHNOLOGY	5
KIVEKAS--O	HELSINKI UNIV TECHNOLOGY	5
LI--ZX	OULU UNIV	5
MANTYJARVI--J	VTT	5
HAARALA--C	UNIV TURKU	4
HALONEN--KAI	HELSINKI UNIV TECHNOLOGY	4
HAMALAINEN--TD	TAMPERE UNIV TECHNOLOGY	4
LAINE--M	UNIV TURKU	4
OLLIKAINEN--J	HELSINKI UNIV TECHNOLOGY	4
RAATIKAINEN--K	UNIV HELSINKI	4
RENFORS--M	TAMPERE UNIV TECHNOLOGY	4
VEIJALAINEN--J	WASEDA UNIV	4

3.1.2. Prolific Journals

Table 2 lists the 21 most prolific journals containing Finnish WNMC research papers. These appear to be applied journals, split between communications and networks. Journals addressing more fundamental issues or topics from other disciplines are not represented. IEEE journals are represented most prominently.

TABLE 2 – MOST PROLIFIC JOURNALS – FINNISH WNMC RESEARCH

JOURNAL	#PAPERS

WIRELESS PERSONAL COMMUNICATIONS	12
IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS	6
MOBILE HUMAN-COMPUTER INTERACTION - MOBILEHCI 2004, PROCEEDINGS	5
IEICE TRANSACTIONS ON COMMUNICATIONS	4
EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS	4
IEE PROCEEDINGS-COMMUNICATIONS	4
BIOELECTROMAGNETICS	4
TELECOMMUNICATIONS AND NETWORKING - ICT 2004	4
IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY	4
IEEE JOURNAL OF SOLID-STATE CIRCUITS	4
IEEE COMMUNICATIONS MAGAZINE	4
ELECTRONICS LETTERS	4
MICROWAVE AND OPTICAL TECHNOLOGY LETTERS	3
IEEE TRANSACTIONS ON COMMUNICATIONS	3
IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION	3
HUMAN-COMPUTER INTERACTION WITH MOBILE DEVICES AND SERVICES	3
COMPUTER COMMUNICATIONS	3
IEEE WIRELESS COMMUNICATIONS	3
SECURITY PROTOCOLS	3
AMBIENT INTELLIGENCE, PROCEEDINGS	3
IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II-ANALOG AND DIGITAL SIGNAL PROCESS	3

3.1.3. Prolific Institutions and Countries

This section identifies the most prolific institutions producing Finnish-authored WNMC papers, and the countries of the most prolific collaborators with Finnish authors of WNMC papers.

Table 3A contains a list of the fifteen most prolific institutions for Finnish-authored WNMC papers, and Table 3B contains a list of the sixteen most prolific countries associated with Finnish-authored WNMC papers. Nokia and Helsinki University of Technology are by far the leaders, with four institutions constituting the second tier: Oulu University, Tempere University of Technology, VTT, University of Helsinki. There are nine universities, four research centers, and two industrial organizations. This distribution reflects an applied research/ technology development emphasis and is consistent with the journal section conclusions. The major collaborators are the advanced Western countries and Japan.

TABLE 3A–MOST PROLIFIC INSTITUTIONS–FINNISH WNMC RESEARCH

INSTITUTION	#PAPERS
NOKIA	51

HELSINKI UNIV TECHNOL	48
UNIV OULU	33
TAMPERE UNIV TECHNOL	31
VTT	19
UNIV HELSINKI	14
UNIV JYVASKYLA	9
UNIV TURKU	9
HELSINKI INST INFORMAT TECHNOL	5
TECH RES CTR FINLAND	4
ABO AKAD UNIV	4
LAPPEENRANTA UNIV TECHNOL	4
SONERA CORP	3
STUK RADIAT & NUCL SAFETY AUTHOR	3
UNITED ARAB EMIRATES UNIV	3

TABLE 3B–MOST PROLIFIC COUNTRIES–FINNISH WNMC RESEARCH

COUNTRY	#PAPERS
FINLAND	210
USA	10
GERMANY	9
ENGLAND	8
SWEDEN	8
JAPAN	7
FRANCE	6
GREECE	5
ITALY	5
CANADA	4
NETHERLANDS	4
U ARAB EMIRATES	4
DENMARK	3
NORWAY	3
AUSTRIA	2
YUGOSLAVIA	2

3.2 Citation Statistics on Authors, Papers, and Journals

The second group of metrics presented is counts of citations to papers published by different entities. While citations are ordinarily used as impact or quality metrics (Garfield, 1985), much caution needs to be exercised in their frequency count interpretation, since there are numerous reasons why authors cite or do not cite particular papers (Kostoff, 1998b; MacRoberts and MacRoberts, 1996). The citation approach recommended is shown in

(Kostoff, 2005b) for a country assessment of Mexico. It will not be repeated here.

Instead, this citation section will present two approaches for assessing Finland's technical thrust areas. One approach is based on document clustering and compares the four themes used for the bibliometrics analysis. They are important research areas for Finland. The other approach is based on phrase frequency analysis and identifies a number of pervasive technical themes that exhibit relatively high activity (numbers of articles published). For both approaches, the Finnish results are compared with those of two similar countries, Norway and Denmark. All three are small Scandanavian countries with similar populations and GDP. All three are viewed as advanced Western nations.

In addition, Appendix 2 contains a slightly different approach for citation comparison. It was the first citation comparison approach tried for the present study. It compares in detail the topics Films and Genes against citation metrics for the countries Mexico and Switzerland.

3.2.1. Document Clustering/ Citation Approach

The general approach was to compare the citation results for papers representative of each of the four research areas (obtained from the document clustering taxonomy) published in a vintage year, focusing on the most cited papers for each country. This metric is selected for research core competencies in particular. While all research in a technical thrust area is important to evaluate, if a thrust area is to be considered as a core research competency, it needs to produce some high impact papers. Especially for high-risk high-payoff research thrusts, a number of low impact papers are acceptable as long as some high impact papers result. The year 1999 was selected as the comparison year. It is of sufficient vintage to allow citations to accumulate, yet sufficiently near-term to ensure relevancy.

All the research articles on the four areas selected from the document clustering-based taxonomy published in 1999 that had at least one author with a Finland or Norway or Denmark address were retrieved, and their citations examined. Then, global citation statistics for each country for each technology were tabulated and compared.

The global comparison component was selected for the following reasons. In evaluating research impact, there are three main criteria to consider:

- 'right job',
- 'job right',
- 'productivity/ progress'.

'Right job' refers to proper selection of the broadest objectives; i.e., is the right study being pursued? 'Job right' refers to selection of the best approaches to solving the problem to reach the desired goal. 'Productivity/ progress' refer to whether anything tangible is being accomplished.

A detailed determination of 'right job' using citation statistics would require clustering the vintage papers thematically, examining citation ranges for each cluster (theme), then assuming that those themes that had the highest citations were the 'hot' research areas. The papers that were in the 'hot' clusters would get high ratings for the 'right job' criterion. The 'job right' rating for any of the papers would be determined by its citation position within any of the clusters. However, for this initial country application of the global comparison, the first two criteria are combined, and the overall citation statistics for all the four selected thrusts' papers will be compared for the three countries.

Table 4 contains the citation comparison results. For each metric, the 'leading' country statistic is shown in **bold**.

TABLE 4 – COMPARISON OF CITATIONS FOR FOUR RESEARCH THRUST AREAS

	FINLAND	NORWAY	DENMARK
POPUL (MILL)	5.2	4.6	**5.4**
GDP (MILL)	151000	**183000**	174000
SCI ART (2004)	7676	5488	**8192**
SCI ART (1999)	7099	4815	**7736**
WIRELESS NETWORKS AND MOBILE COMMUNICATIONS (WNMC)			
ART (2004)	**106**	18	14
ART (1999)	**32**	5	6
MED-20%	35	**49**	28
MED-ALL	2	**18**	9
#>40	**3**	1	0
SIGNAL PROCESSING (SIGPR)			
ART (2004)	**87**	50	59

ART (1999)	**67**	33	58
MED-20%	17	17	**28**
MED-ALL	3	4	**6**
#>20	6	3	**10**
MATERIALS SCIENCE AND ENGINEERING (MATLS)			
ART (2004)	**541**	259	481
ART (1999)	**396**	213	360
MED-5%	44	34	**76**
MED-ALL	6	5	**7**
#>40	14	3	**32**
CHEMISTRY (CHEM)			
ART (2004)	1206	773	**1363**
ART (1999)	1130	700	**1268**
MED-5%	70	67	**76**
MED-ALL	9	8	**12**
#>90	13	7	**22**

CODE

SCI ART (2004) - NUMBER OF RESEARCH ARTICLES PUBLISHED IN SCI IN 2004
SCI ART (1999) - NUMBER OF RESEARCH ARTICLES PUBLISHED IN SCI IN 1999
MED-YY% - MEDIAN CITATIONS OF THE TOP YY% OF ALL CITED ARTICLES
MED-ALL - MEDIAN CITATIONS OF ALL 1999 CITED ARTICLES
#>ZZ - NUMBER OF ARTICLES WITH CITATIONS GREATER THAN ZZ

The first four rows in Table four contain macro-level statistics for the three countries. For 2005 (shown), the populations are within fifteen percent of each other, and the GDPs are within twenty percent. Normalization of the citation statistics to either population or GDP will not be used, since the numbers are relatively close, and it is not clear which number would be more appropriate for normalization.

The total number of research articles in the SCI for both 1999 and 2004 (rows three and four in Table 4) are quite similar for Finland and Denmark and noticeably less for Norway. Whether this reflects reduced Norwegian investment in research, or reduced research productivity, or both, (or neither) is unclear at this point.

3.2.1.1. Wireless Networks and Mobile Communications (WNMC)

WNMC was defined by the query shown previously, which was used to retrieve articles from the SCI published in 1999. As shown in Table 4, the number of Finnish articles in WNMC published in both 1999 and 2004 is almost an order of magnitude larger than that from Norway or Denmark.

This investment reflects an obvious strategic priority for Finland and is one characteristic of a true core competency.

Because of the small number of data points for Norway and Denmark for the citation comparison vintage year (1999), the citation comparison statistics have limited validity. For example, the median of the top twenty percent of articles for Norway and Denmark reflected one article for each country. In addition, many of the Norway and Denmark articles did not focus on development of the wireless/ mobile technology, but rather on consequences such as health effects of microwave radiation. Finland did have three highly cited papers (>40), but the relative significance of this number could not be ascertained due to the poor statistics for the other two countries.

To place this result in a larger context, the citation performance of the USA in WNMC was examined for articles published in 1999 and compared to that of Finland.

- There were 943 USA articles published in 1999 (compared to Finland's 106).
- The median of the top twenty percent USA article citations was 32 (compared to Finland's 35).
- The median of all the USA article citations was 5 (compared to Finland's 2).
- Finally, the number of USA articles with forty or more citations was 69 (compared to Finland's 3).

Thus, in relative terms Finland's research output in WNMC reflects a higher priority than the USA's, but in absolute terms, the USA is dominant by a large margin.

3.2.1.2. Signal Processing (SIGPR)

SIGPR was defined by the following query:

SIGNAL PROCESSING OR DATA FORMAT CONVERTER* OR BANDPASS SAMPL* OR ((SIGNAL OR DIGITAL) AND (FILTER* OR CODE OR CODING OR TRANSMIT* OR AUDIO OR VIDEO OR SPEECH OR IMAGE OR COMMUNICATION OR GEOPHYSIC* OR ACOUSTIC* OR SONAR OR RADAR OR MUSIC* OR ANALOG CONVERTER* OR WAVEGUIDE OR WAVE GUIDE OR RECEIVER*))

Since 1999, Norway has increased its output substantially in SIGPR, while Finland's output has increased moderately, and Denmark's has remained the same. Denmark exceeds Finland and Norway noticeably in all the citation metrics used. Most importantly, Denmark has a higher fraction of the most highly cited articles. *Thus, while SIGPR may be a core Finnish research area due to volume of activity, it does not offer a core research impact advantage over Denmark*.

3.2.1.3. Materials Science and Engineering (MATLS)

Materials research was defined by the following query:

(NANO* NOT (NANOSECOND* OR NANOLITER* OR NANOMOLAR* OR NANOBACTERIA OR NANO3 OR NANOFILTRATION OR NANOCOLLOID*)) OR SMART MATERIAL* OR METALS OR METALLIC OR ALLOY* OR CORROSI* OR BIOCHEMICAL MATERIAL* OR THIN FILM* OR CERAMIC* OR (MATERIAL* SAME (COPOLYMER* OR POLYMER* OR FILM*)) OR (COMPOSITE* SAME (REINFORCE* OR BOND* OR RESIN* OR STRENGTH OR FLEXURE OR GLASS OR FRACTURE OR FIBER OR SUBSTRATE OR PROPERTIES OR MATERIAL*)) OR ATOMIC FORCE MICROSCOPY OR SCANNING ELECTRON MICROSCOPY OR SCANNING TUNNELING MICROSCOPY OR CHEMICAL VAPOR DEPOSITION OR STEEL OR STEELS OR SLAG OR TENSILE TEST* OR X-RAY DIFFRACTION OR ATOMIC LAYER OR SURFACE ROUGHNESS

Finland and Denmark are clear leaders over Norway in numbers of research articles published, with Finland holding a slight edge. Both Finland and Denmark have increased publication productivity by about a third since 1999, while Norway's increase has been about half that amount. As in the SIGPR thrust, Denmark outperforms Finland in all the citation metrics used, most importantly in the number of highly cited papers. *Again, as in the SIGPR thrust, Materials research may be a core Finnish research area due to volume of activity, but it certainly does not offer a core research impact advantage over Denmark.*

3.2.1.4. Chemistry (CHEM)

Chemistry research was defined by the following query:

CHEMISTRY OR BIOCHEMISTRY OR CHEMICAL OR BIOCHEMICAL OR CATALYST* OR CATALYTIC OR CATALYSIS OR (REACTION* AND (ENANTIOSELECT* OR KINETIC* OR HYDROGEN OR OXYGEN OR ACID OR ACIDS OR SOLVENT* OR SOLUTION* OR OXIDATION OR COMPLEX*)) OR (COMPOUND* AND (SYNTHESIS OR DERIVATIVE* OR ACID* OR BOND* OR MOLECUL* OR CLEAVAGE*)) OR (NMR AND (METHYL OR RING* OR COMPOUND* OR STRUCTURE* OR PHENYL OR REACTION* OR ISOMER*)) OR (COMPLEXES AND (LIGAND* OR BOND* OR ION* OR CATION* OR ATOM* OR STRUCTURE* OR ANION*)) OR ((POLYMER* OR CO-POLYMER*) AND (CHAIN* OR BLOCK* OR SYNTHESIS OR COMPLEX*)) OR (SOLUTION* AND (ACID* OR SODIUM OR AQUEOUS OR ION EXCHANGE)) OR (SULFURIC ACID AND WATER) OR LIQUID CHROMATOGRAPH* OR MASS SPECTROMET* OR POLYCYCLIC AROMATIC HYDROCARBON* OR RADICAL SCAVENG*

Again, Finland and Denmark are the clear leaders in numbers of CHEM publications, and are relatively close to each other. *Again, Denmark is the clear leader in all the citation metrics used, with moderately less of the commanding lead shown in Materials research.*

Three caveats are in order here. The first concerns the relation between citations, research impact, and core competencies. Poor citation performance can reflect:

- Poor intrinsic quality, and/ or
- Low circulation journals, and/ or
- Low research activity in field, and or
- More applied focus, reducing the number of extra-discipline researchers available to cite, and/ or
- other characteristics.

In particular, a country could produce high quality but very applied research in a technical area. The area's output might receive low citations in aggregate but still be a national core competence. In order to distinguish the reasons for poor citations exhibited by different countries in different

sections of Table 4, one would need to read substantial numbers of research articles produced by the countries in the field of interest, and analyze them for intrinsic quality and level of development. That was beyond the scope of the present study.

The second caveat concerns the aggregation level of the technologies examined. The conclusions above about dominance are for the thrust areas defined by the relatively broad queries shown. *Even though one country may dominate on the thrust area in aggregate (as defined by the total query), another country could conceiveably lead in one of the sub-thrust areas*.

The third caveat concerns the selection of phrases. The selection was based on emphasis in the Finland research output database. A priori, one would expect these to be Finnish research priority areas, and would expect Finland's performance relative to other countries to be somewhat higher than average. Any poor relative performance by Finland in these Finnish thrust areas should be cause for concern.

In the above four thrust areas analyzed, Finland has two characteristics of a true core competency: an identifiable cohesive thrust and identifiable critical mass (numbers of publications, relative to Norway and Denmark). However, for the latter three technical thrust areas, it did not have the aggregate citation impact of Denmark. The differences in relative publication and citation attributes across the four technical areas for the three countries show the necessity for comparing countries at the critical technology level rather than at the aggregate national level (King, 2004).

3.2.2. Phrase Frequency/ Citation Approach

A phrase frequency analysis was performed on the contents of the Abstracts of all the Finland records retrieved. The highest frequency technical phrases were identified. Seventeen of these phrases were selected from the areas of Physical, Environmental, and Life Sciences, with emphasis on phrases different from those used for the core competency queries. No high frequency phrases of adequate specificity could be identified from the Engineering Sciences area. Each phrase (in some cases, combinations of similar phrases) was entered into the SCI search engine, and records were retrieved for Finland, Norway, and Denmark for 1999. Citation indicators were applied to each phrase's retrievals, and the results compared.

Table 5 contains the comparisons. The first seven phrases/ phrase combinations relate to Life Sciences research (Genetics; Cancer; Diabetes; Cardiovascular; Immunity; Infection; Proteins), the next three relate to Environmental Sciences (Soils; Forestry; Climate), and the last seven relate generally to Physical Sciences (Electron Microscopy; Astronomy; Optics; Mass Spectrometry; Molecules; Atoms; Superconductivity), although some (e.g., Molecules, Atoms) are sufficiently general to encompass other disciplines as well.

TABLE 5 – CITATION ANALYSIS OF SELECTED HIGH FREQUENCY TECHNICAL PHRASES/ PHRASE COMBINATIONS

	FINLAND	NORWAY	DENMARK
GENE OR GENES OR GENETICS OR (GENETIC NOT (GENETIC ALGORITHM* OR GENETIC PROGRAM*))			
ARTICLE-1999	780	419	**822**
MED CITES - ALL	**16**	12	14
MED CIT-TOP 5%	**122**	90	99
MED CIT-TOP 10	**207**	133	180
#>100	**26**	8	20
CANCER* OR CARCINOMA* OR TUMOR* OR MALIGNAN*			
ARTICLE-1999	**590**	407	455
MED CITES - ALL	**15**	10	13
MED CIT-TOP 5%	**155**	99	107
MED CIT-TOP 10	**227**	133	178
#>100	**21**	10	15
DIABET* OR INSULIN*			
ARTICLE-1999	199	63	**349**
MED CITES - ALL	**16**	14	12
MED CIT-TOP 10%	**119**	106	69
MED CIT-TOP 10	**209**	84	108
#>100	**13**	3	7
CARDIOVASCULAR OR MYOCARDIAL OR BLOOD PRESSURE OR HYPERTENSION			
ARTICLE-1999	**324**	166	287
MED CITES - ALL	**12**	9	9
MED CIT-TOP 10%	**102**	68	74
MED CIT-TOP 10	**280**	83	119
#>100	**17**	3	9
ANTIBOD* OR ANTIGEN* OR IMMUN*			
ARTICLE-1999	**806**	453	764
MED CITES - ALL	**13**	12	12
MED CIT-TOP 5%	**120**	98	100
MED CIT-TOP 10	**275**	165	208

#>100	**23**	11	20
INFECT*			
ARTICLE-1999	357	207	**435**
MED CITES - ALL	10	10	**11**
MED CIT-TOP 5%	95	**107**	103
MED CIT-TOP 10	**193**	107	145
#>100	8	7	**12**
PROTEIN*			
ARTICLE-1999	880	529	**1035**
MED CITES - ALL	**15**	11	13
MED CIT-TOP 10%	**81**	64	**81**
MED CIT-TOP 10	**286**	134	217
#>100	25	9	**34**
SOIL* OR SEDIMENT*			
ARTICLE-1999	200	250	**372**
MED CITES - ALL	9	7	**11**
MED CIT-TOP 10%	35	36	**39**
MED CIT-TOP 10	43	55	**63**
#>40	8	9	**18**
FOREST* OR WOOD			
ARTICLE-1999	**254**	95	97
MED CITES - ALL	8	9	**10**
MED CIT-TOP 10%	35	35	**61**
MED CIT-TOP 10	45	35	**61**
#>40	**9**	4	7
ATMOSPHERE* OR CLIMAT*			
ARTICLE-1999	126	155	**165**
MED CITES - ALL	7	10	**11**
MED CIT-TOP 10%	**68**	67	58
MED CIT-TOP 10	75	**82**	64
#>40	13	**18**	15
ELECTRON MICROSCOP*			
ARTICLE-1999	**93**	86	90
MED CITES - ALL	7	6	**8**
MED CIT-TOP 10%	**67**	62	46
MED CIT-TOP 10	**67**	60	45
#>50	**6**	**6**	4
(TS=(GALAX* OR GALACTIC OR INTERSTELLAR) OR SO=(ASTRON* OR ASTROPHYS*)) AND CU=FINLAND AND PY=1999			
ARTICLE-1999	62	19	**84**
MED CITES - ALL	11	6	**15**
MED CIT-TOP 10%	27	32	**55**
MED CIT-TOP 10	25	17	**53**
#>30	2	1	**15**

OPTIC*			
ARTICLE-1999	174	87	**241**
MED CITES - ALL	5	7	**10**
MED CIT-TOP 10%	27	34	**68**
MED CIT-TOP 10	33	33	**92**
#>30	7	6	**43**
MASS SPECTROMET*			
ARTICLE-1999	79	50	**119**
MED CITES - ALL	9	9	**14**
MED CIT-TOP 10%	57	39	**105**
MED CIT-TOP 10	55	29	**124**
#>40	9	2	**21**
MOLECUL*			
ARTICLE-1999	633	385	**740**
MED CITES - ALL	11	9	**13**
MED CIT-TOP 5%	91	79	**101**
MED CIT-TOP 10	**207**	97	176
#>100	**12**	4	**20**
ATOM*			
ARTICLE-1999	191	114	**260**
MED CITES - ALL	7	7	**10**
MED CIT-TOP 10%	37	38	**68**
MED CIT-TOP 10	50	39	**110**
#>40	7	4	**38**
SUPERCONDUCT*			
ARTICLE-1999	64	22	**64**
MED CITES - ALL	5	10	**5**
MED CIT-TOP 10%	30	24	**30**
MED CIT-TOP 10	26	16	**26**
#>30	3	0	**3**

CODE

MED CITE-ALL - MEDIAN CITATIONS OF ALL 1999 CITED ARTICLES
MED CIT-TOP X% - MEDIAN OF TOP X% OF ALL 1999 CITED ARTICLES
MED CIT-TOP 10 - MEDIAN OF TOP 10 OF ALL 1999 CITED ARTICLES
#>Y - NUMBER OF 1999 ARTICLES RECEIVING GREATER THAN Y CITATIONS

The phrase frequency approach provides a somewhat different, although complementary, perspective on the three countries' performance from the document clustering/ citation approach. The metrics for the 'winning' country for each phrase/ phrase combination are bolded.

The phrase frequency approach shows:

- Finland to be the clear winner in most of the citation metrics in most of the Life Sciences themes,
- Denmark to be the winner in most Environmental Sciences themes,
- Denmark to be a clear winner in four of the seven Physical Sciences themes,
- Finland to be a modest winner in one Physical Sciences theme (Electron Microscopy), and
- Finland and Denmark to be essentially tied in three Physical Sciences themes.

The caveats expressed previously about the interpretation of citations and thematic aggregation apply to these results as well.

The real surprise in these results is the citation performance of Norway. While Norway has the smallest population of the three countries, it has the largest GDP. *In most cases (but not all), Norway had the lowest research activity, and in almost all cases, had the lowest citation impact*. It should be re-stated that the phrases selected for the present country comparison analysis, whether from the document clustering-based queries or the phrase frequency-based queries, were based on emphasis in the Finland data base. Whether Norway would perform better if phrases were selected from an analysis of its database is unknown. As will be shown in the next paragraph, its gross citation numbers do not offer much encouragement, but as this study has emphasized repeatedly, the dis-aggregated critical technologies are of most interest.

The overall citation statistics of the SCI articles from 1999 for the three countries were generated and are shown in Table 6. The purpose of this table is to identify the effect of collaborating countries on the research impact of articles with Finnish authors. There are four separate groups of results displayed. The first group, Total Country Results, is based on all articles from either 2004 (first row) or 1999 (all other rows) that had at least one author with the address of the country heading the column (i.e., Finland or Norway or Denmark). The second group, Country Only (No Other Countries), is based on all articles from 1999 that had at least one author with the address of the country heading the column, excluding all authors with an address from any of Finland's 25 top country collaborators. The

purpose of this group is to identify citation characteristics of papers from only the country heading the column (Finalnd, Norway, or Denmark).

TABLE 6 – MACRO COUNTRY RESULTS

	FINLAND	NORWAY	DENMARK
TOTAL COUNTRY RESULTS			
SCI ART (2004)	7676	5488	**8192**
SCI ART (1999)	7099	4815	**7736**
MED CIT-ALL	7	7	**8**
MED CIT-TOP 1%	**169**	134	167
MED CIT-TOP 50	**207**	127	196
#>300	10	4	**12**
COUNTRY ONLY (NO OTHER COUNTRIES)			
SCI ART-1999	**4328**	2814	4062
MED CIT-ALL	6	5	**7**
MED CIT-TOP 1%	93	82	**108**
MED CIT-TOP 50	90	69	**103**
#>300	1	0	**2**
COUNTRY AND USA ONLY (NO OTHER COUNTRIES)			
SCI ART-1999	499	344	**555**
MED CIT-ALL	**13**	8	**13**
MED CIT-TOP 1%	**371**	135	161
MED CIT-TOP 50	**83**	44	74
#>300	**3**	0	1
COUNTRY AND USA			
SCI ART-1999	875	614	**1012**
MED CIT-ALL	**14**	11	**14**
MED CIT-TOP 1%	**367**	306	343
MED CIT-TOP 50	**117**	88	108
#>300	5	3	**6**

CODE

MED CITE-ALL - MEDIAN CITATIONS OF ALL 1999 CITED ARTICLES
MED CIT-TOP 1% - MEDIAN OF TOP 1% OF ALL 1999 CITED ARTICLES
MED CIT-TOP 50 - MEDIAN OF TOP 50 OF ALL 1999 CITED ARTICLES
#>300 - NUMBER OF 1999 ARTICLES RECEIVING GREATER THAN 300 CITATIONS

The third group, Country and USA Only (No Other Countries) is based on all articles from 1999 that had at least one author with the address of the country heading the column and at least one author with a USA address, excluding all other authors with an address from any of Finland's 25 top country collaborators. The purpose of this group is to identify the impact of

adding only USA co-authors to those of the country heading the column. The final group, Country and USA, is based on all articles from 1999 that had at least one author from the USA and one author from the country heading the column. Papers with authors from other countries are not excluded, as they were in the third group.

The summary findings are:

- In the first group, Total Country Results, the gross citation impacts of Finland and Denmark are quite similar and substantially larger than that of Norway.
- In the second group, Country Only, Denmark has modestly better citation performance than Finland, which in turn has modestly better citation performance than Norway.
- In the third group, Country and USA Only (No Other Countries), Finland and the USA outperform Denmark and the USA. Both groups substantially outperform Norway and the USA.
- In the fourth group, Country and USA, multi-country papers including Finland and USA very slightly outperform multi-country papers including Denmark and USA, which in turn moderately outperform multi-country papers including Norway and the USA.
- Adding the USA only to Finland/ Norway/ Denmark papers increases the citation performance substantially, and adding more countries enhances performance modestly for Finland, somewhat more for Denmark, and substantially for Norway.

3.3. Taxonomies

Based on the complete set of 12900 retrieved papers, statistical document clustering was performed. Document clustering is the grouping of similar documents into thematic categories. Different approaches exist (Cutting et al, 1992; Guha et al, 1998; Karypis et al, 1999; Rasmussen, 1992; Willet, 1998; Zamir and Etzioni, 1998). The specific document clustering approach used for the present study is based on a partitional clustering algorithm (Karypis, 2004) contained within a software package named CLUTO. Most of CLUTO's clustering algorithms treat the clustering problem as an optimization process that seeks to maximize or minimize a particular clustering criterion function defined either globally or locally over the entire clustering solution space. CLUTO uses a randomized incremental

optimization algorithm that is greedy in nature, and has low computational requirements.

Specifically, CLUTO implements various algorithms for clustering low- and high-dimensional datasets and for analyzing the characteristics of the various clusters. CLUTO implements three different classes of clustering algorithms that can operate either directly in the object's feature space or in the object's similarity space. The clustering algorithms provided by CLUTO are based on the partitional, agglomerative, and graph-partitioning paradigms. CLUTO's partitional and agglomerative algorithms are able to find clusters that are primarily globular, whereas its graph-partitioning and some of its agglomerative algorithms are capable of finding transitive clusters.

In this study, documents were clustered using the partitional clustering algorithms provided by CLUTO. Partitional clustering algorithms find the clusters by partitioning the entire document collection into a predetermined number of disjoint sets, each corresponding to a single cluster. This partitioning is achieved by treating the clustering process as an optimization procedure that tries to create high quality clusters according to a particular function that reflects the underlying definition of the “goodness” of the clusters. This function is referred to as the *clustering criterion function.* CLUTO implements seven such criterion functions that measure various aspects of intra-cluster similarity, inter-cluster dissimilarity, and their combinations, and have been shown to produce high-quality clusters in low- and high-dimensional datasets.

CLUTO uses two different methods for computing the partitioning clustering solution. The first method computes a *k*-way clustering solution via a sequence of repeated bisections, whereas the second method computes the solution directly (in a fashion similar to traditional *K*-means-based algorithms). These methods are often referred to as *repeated bisecting* and *direct k*-way clustering, respectively. CLUTO computes a direct *k*-way clustering as follows. Initially, a set of *k* objects is selected from the datasets to act as the *seeds* of the *k* clusters. Then, the similarity of each object to these *k* seeds is computed, and the object is assigned to the cluster corresponding to its most similar seed. This forms the initial *k*-way clustering. This clustering is then repeatedly refined so that it optimizes a desired clustering criterion function. This optimization is performed using a randomized incremental optimization algorithm that is greedy in nature, has low computational requirements, and produces high-quality solutions (Zhao

and Karypis, 2004). A *k*-way partitioning via repeated bisections is obtained by recursively applying the above algorithm to compute 2-way clustering (*i.e.*, bisections). Initially, the objects are partitioned into two clusters, then one of these clusters is selected and is further bisected, and so on. This process continues *k* - 1 times, leading to *k* clusters. Each of these bisections is performed so that the resulting two-way clustering solution optimizes a particular criterion function.

The actual documents were represented using the widely-used vector-space model in which the various terms present in the documents were used to define a high-dimensional space and each document was considered to be a vector in that space. However, unlike the traditional vector-space representation, which relies entirely on single terms, all consecutive two- and three-word combinations were taken into account, resulting in a representation that is capable of capturing the phrases commonly occurring in the documents. In addition, Porter's stemming algorithm was used to pre-process the various terms of each document prior to obtaining their vector-space representation. The weight of each dimension was computed using the TF-IDF model in which terms that occur many times within a document are given higher weight (TF) and terms that occur across many documents were given lower weight (IDF). The similarity between two documents was measured using the cosine of their corresponding document vectors.

3.3.1. Partitional Clustering Results

In partitional clustering, the number of clusters desired is input, and all documents in the database are included in those clusters. The 200 clusters selected for the present study were aggregated into a hierarchical taxonomy using a hierarchical tree generated by the CLUTO software. The taxonomy is shown in Figure 1. The 200 individual clusters are shown in detail in Appendix 3.

The taxonomy and its component category levels are generated as follows. The hierarchical tree for the total database of retrieved articles is generated by the clustering software. Each node in the hierarchical tree has a number of associated features: number of articles in node, most frequent weighted and unweighted phrases. A separate output contains the article Abstracts for each of the 200 elemental clusters.

APPROACH AND RESULTS

To generate a theme for each node, as shown on Figure 1, the analyst examines the most frequent weighted and unweighted phrases, reads a sample of the Abstracts for the node being examined, then specifies the theme that represents the phrases and their contexts best.

The categories in the taxonomy levels, and the number of documents in each category, are described as follows.

FIGURE 1 – PARTITIONAL DOCUMENT CLUSTERING TAXONOMY

Level 1	Level 2	Level 3	Level 4
Physical, Informational, Environmental and Economic Sciences (7158)	Physics, Information Science, and Economic Science (3579)	Economic Studies, Information Technology, Theoretical Mathematics (802)	Economics / Sociology / Economic Models (498)
			Information Technology / Communication Systems (304)
		Theoretical Mathematics and Applied Physics (2777)	Theoretical Mathematics / Signal Processing (2037)
			Applied Physics (740)
	Material and Chemical Sciences, Environment, and Ecology (3579)	Material and Chemical Science, Organic (1555)	Physical Chemistry / Metallurgy / Material Science (559)
			Organic Chemistry (996)
		Environmental and Ecological Studies (2024)	Environmental Studies (1291)
			Ecological Studies (733)
Biochemistry and Biomedical (5415)	Biochemistry, Biology, and Medicine (2110)	Biochemistry (1597)	Laboratory Biological Trials (744)
			Biomolecular Complexes (853)
		Genetic Science and Medicine (513)	Genetics of Cancerous Cells (315)
			Cancer Risk Factors and Detection (198)
	Clinical Medicine (3305)	Treatment of Medical Conditions (1274)	Surgical Treatment, Medicines (659)
			Medicines and Medical Treatment Equipment (615)
		Lifestyle, Pregnancy, and Neuropsychological (2031)	Healthcare, Long Term Medical Conditions, Mental Disorders (1374)
			Pregnancy and Neuropsychological (657)

On Figure 1, the columns represent the taxonomy levels. There are four levels depicted in this taxonomy. The highest level (two categories) is the first column, and the lowest level shown (16 categories) is the last column. The numbers in parentheses represent the numbers of records assigned to the category.

The first level has two categories: Physical, Informational, Environmental and Economic Sciences (7158) and Biochemistry and Biomedical (5415). In Figure 1, the second taxonomy level is generated by sub-dividing each first level category by two. Physical et al divides into Physics, Information Science, and Economic Science (3579) and Material and Chemical Sciences (3579), while Biochemistry and Biomedical divides into Biochemistry, Biology, and Medicine (2110) and Clinical Medicine (3305).

The fourth level on Figure 1 can be viewed as a flat taxonomy representation of Finland's research structure. The sixteen categories comprising the flat taxonomy will be described in detail in the following analysis. The thrusts (individual cluster themes from the 200 elemental clusters) under each fourth level category are shown in bullets. Numbers of records under each thrust are shown in parentheses preceding the definition. The categories below are arranged in order of their appearance on Figure 1, starting from the top of level 4.

Economics / Sociology / Economic Models

- (42) factors effecting the operation of Finnish businesses and companies, particularly new markets, establishment of multi-national corporations (MNCs), and growing environmental complexities.
- (58) the state of the economy.
- (65) economic trends and economic strategies.
- (90) social, economic, and political aspects of Finland.
- (108) sociological concepts and theories.
- (41) the use of information technology in the educational process.
- (45) decision making processes (such as analytic hierarchy process (ahp)) and decision support systems.
- (49) core recovery accident management research efforts / programs.

Information Technology / Communication Systems

- (90) software engineering, and software development methods / processes.
- (68) (video) games designed for wireless networks and mobile communication systems.
- (67) mobile computing, communication system, and fusing sensory applications.
- (79) wireless networks and mobile communication devices, with emphasis on protocols and security.

Theoretical Mathematics / Signal Processing

- (27) mathematical algorithms and linear codes.
- (65) communication technologies, particularly cellular communication network receivers and decoders, channel estimation processes, and channel interference problems.
- (131) algorithms for various applications.
- (58) optimization algorithms, and signal processing methods such as finite impulse response (FIR).
- (66) the finite automata model of computation, and the theory of formal languages.
- (43) mathematical models.
- (81) nonlinear mathematics and sobolev spaces.
- (83) solutions to differential equations and their use in practical applications.
- (63) the theory of phase transitions and critical phenomena.
- (61) nonlinear dynamic models, with emphasis on noise and speech (recognition).
- (109) the use of models to simulate biological and physical phenomena.
- (102) methods for analyzing data sets.
- (56) visualization algorithms such as self organizing map.
- (72) the physics of (electronic) imaging systems.
- (81) signal processing, signal filtering techniques and noise reduction.
- (83) communications equipment (particularly antennas), antenna calibration, and communication frequencies.
- (63) laser technologies, particularly laser welding, and vertical cavity surface emitting laser for applications such as high speed optical

interconnects, parallel data links, optical recording, 2-D scanning, and optical signal processing.

- (79) use of laser beam devices such as diffractive optical elements.
- (60) astronomy, particularly luminous stars and active galactic nuclei.
- (46) asteroid orbit and topography, and lunar orbit and topograpy.
- (46) geological instrumentation, and velocity measurement techniques using radar.
- (59) the mechanics, manufacture, and use of structural and nonstructural materials in a variety of applications.
- (32) the use of finite difference time domain (fdtd) methods to model / analyze mechanical properties of materials and macromolecular structures and articular cartilage.
- (88) processes / reactions that create radiative heat transfer as a byproduct, and single-scattering modeling for radiative flux and radiance calculations.
- (59) waves and spatial dispersion.
- (48) solar activity and interplanetary magnetic fields.
- (53) the ionosphere, interplanetary magnetic fields and solar winds, particularly aurora borealis phenomena.
- (111) magnetic fields and materials, particularly, magnetic shape memory alloys.
- (44) the propagation and retrieval of cross-spectral density.
- (43) methods and processes for measuring magnetic and electric fields.

Applied Physics

- (55) plasma physics and gyrotron systems.
- (51) gaseous superfluid phases, vortices, and bose einstein condensates.
- (54) josephson junctions, joshephson coupling, tunneling and superconducting metals.
- (79) physics used for calculating the electronic structure, properties of matter with methods such as spin deinsity functional theory, and analyzing the spectra of quantum dots.
- (77) molecular dynamic simulations, and use of monte carlo methods for evaluating nanocluster.
- (98) Focused molecular quantum mecanics, particularly on potential energy surface (PES).

- (41) experimental tools / methods for spectroscopy, particularly Auger electron spectroscopy.
- (66) the approach for describing the ground state properties of metals, semiconductors, and insulators.
- (69) experimental and theoretical physics, particularly gamma-ray decay.
- (76) theoretical models used in particle physics, with emphasis on hadrons, quark masses, and bosons.

Physical Chemistry / Metallurgy / Material Science

- (30) experimental physics in material research, and analysis of vacancies / defects in material crystal lattices, with emphasis on positron annihilation.
- (54) ion beam analysis techniques for quantitative analysis of light elements in solids, with emphasis on the luminescent properties of annealed or doped materials.
- (81) methods, such as atomic force microscopy, for determining material surface structures at the atomic level.
- (40) metallurgy, memory shaped alloys, and nanomaterials.
- (44) metallurgy, with emphasis on gas bubbles resulting during the metal separation / slag removal process, and effects of alloying elements in the corrosion resistance of steels.
- (34) (fiber reinforced) composite materials.
- (37) biochemical reactions, particularly of metal binding sites.
- (72) substrates and coatings, particularly silica, and their use in various applications.
- (104) deposition and growth of thin film materials for microelectronics applications.
- (48) material science of thin films, microelectronics, processes such as atomic layer deposition, and ink resins for printing.

Organic Chemistry

- (92) the use of selective catalytic hydrogen in chemical reactions.
- (79) organic reactions that allow chemical transformations.
- (54) chemical reactions, synthesis of compounds, and analyzing the molecular mechanics and dynamics of various substances.
- (61) nuclear magnetic resonance (NMR) spectroscopy.
- (107) chemical reactions of various compounds, particularly ligand exchange (substitution) reactions involving complex metal ions.

- (35) various material microstructures and crystal lattices, with an emphasis unit cell parameters.
- (42) synthesis of polymer / copolymer materials.
- (42) chemical reaction of various compounds, especially reactions involving membranes and lipid bilayers.
- (54) chemical reactions involving ionic solutions / sodium chloride solutions.
- (64) water purification and water treatment, with emphasis on the effects of sulfuric acid on water.
- (98) analytical methods of chemistry such as liquid chromatography and mass spectrometry.
- (44) biochemistry, with emphasis on the extraction of polycyclic aromatic hydrocarbon (PAH) compounds.
- (39) oxidation reactions and iron-oxidation.
- (48) wastewater treatment, especially sludge removal.
- (58) the conversion of wood into pulp, chips, and other products, as well as the extraction of lignin from wood.
- (79) pulp and paper production technologies.

Environmental Studies

- (78) issues related to nanoparticles (size, stability, etc.) in various applications, such as drugs and structural materials.
- (70) particle nucleation and aerosols.
- (39) air quality and how it is effected by construction, power generation, particularly fluidized bed combustion, coal ash / dust and fuel emissions.
- (44) heat (mass) transfer.
- (62) the thermodynamics of phase transitions in various chemical reactions.
- (36) mining, granite rocks, and platinum group elements (PGE).
- (63) the effects of regional and global climate changes on Finland.
- (40) ice core research for identifying history and future climatic and environmental patterns.
- (88) the nutrient input, salinity, and other biological properties of the Baltic sea.
- (62) water contamination, water treatment.
- (37) agriculture, and the analysis of sediments and nutrients in water..
- (63) the effects of emissions on the Finnish ecosystem, especially lakes and other bodies of water.

- (53) the effect of CH4 and CO2 emissions on the environment.
- (32) the production of milk / dairy products.
- (145) forestry and the effects of soil properties in the growth of particular tree species.
- (43) forestry, wood, and wood-based products, with an emphasis on carbon sinks and their effects on climate change.
- (87) forestry, particularly forest / ecological planning, species diversification, and landscape management.
- (76) forestry with particular emphasis on diameter distribution of stock, and models for (tree / forest) growth simulation.
- (50) forestry, particularly the scots pine.
- (34) the effect of ozone exposure on the ecosystem.
- (89) particular species of trees, such as the birch betula pendula, and leaf damage due to (insect) herbivores.

Ecological Studies

- (70) various species (particularly birds), their reproductive habits and population trends.
- (52) genetic variability of species and breeding patterns.
- (63) reproductive issues for human, animals, and insects.
- (17) the effect of microorganisms (e.g., argulus coregoni, Acartia bifilosa) on the reproductive behavior of fish.
- (27) the effect of water chemistries on fish populations, especially salmon.
- (53) fishing industry and fish farms, with emphasis on stock, whitefish, and perch..
- (46) various predatory species.
- (43) wildlife habitats and migration patterns, particularly of birds.
- (144) forest ecology research, particularly quantity and richness of various plant species.
- (40) ecological species, particularly host plants, parasites, larvae, fish.
- (45) the polymerase chain reaction for generating copies of fragmented DNA.
- (42) treatment for various viral and bacterial diseases.
- (91) various bacteria strains, especially isolated gene sequences.

Laboratory Biological Trials

- (47) the genetic makeup of various viruses.

- (72) genetic structure of proteins and amino acids.
- (23) heat shock proteins and their role in cancer, especially nephrin and phosphor-lipid transfer proteins (PLTP).
- (54) binding proteins, biding domains, and their biochemical makeup.
- (37) biochemical reactions, particularly of metal binding sites.
- (43) biological synthesis and evaluation of peptide proteins for applications such as hormonal regulation and antibiotic activities.
- (66) biomedical research on peptide binding.
- (110) biochemistry, particularly of (antagonist) receptors and their biding ability.
- (80) genetic research, testing the influence of ethanol, dopamine, and alcohol, on biomedical models such as winstar rats and sprag dawley rats.
- (70) inhibitors to platelet derived growth factor (pdgf).
- (44) the effect of particular proteins in the normal development of the brain., particularly Brain-Derived Neurotrophic Factor (BDNF) and neurons known to develop neurofibrillary tangles (NFTs).
- (98) various clinical trials using transgenic mice, collagen (from calcified and non-calcified cartilage).

Biomolecular Complexes

- (67) the biological process cell adhesion and the extracellular matrix, with emphasis on integrins.
- (54) apoptosis / cell death, and caspase enzymes.
- (159) research utilizing cell lines, with emphasis on (DNA / RNA) expressions.
- (83) mast cells, cell cultures, and epithelial cells.
- (42) the effect of the vascular endothelial growth factor on the blood and lymph systems.
- (118) tissue growth and (cancerous) cell reproduction, with emphasis on (DNA / RNA) expressions.
- (46) the body's immune response to various stimulants, with emphasis on (DNA / RNA) expressions.
- (72) various methods for analyzing RNA and DNA, with emphasis on RNA/DNA expressions.
- (120) proteins that regulate various cell activities.
- (92) root causes of particular genetic disorders.

Genetics of Cancerous Cells

- (72) genome wide scans / genome wide linkages to various health conditions, with emphasis on chromosomes, gene markers, and gene families.
- (105) genetic variations such as single nucleotide polymorphisms.
- (138) genetic makeup of cancerous cells and their hereditary nature, with particular emphasis on gene mutations and gene families.

Cancer Risk Factors and Detection

- (55) various matrix metalloproteinases (MMPs) enzymes and their implication on cancer growth.
- (94) cancer, particularly breast cancer, clinical trials identifying common risk factors.
- (49) measurement techniques for (early) detection of cancer.

Surgical Treatment, Medicines

- (53) cancer patients, particularly lung cancer survivors, and the treatments they took including chemotherapy treatments such as docetaxel.
- (42) Focused patient survival of organ transplant surgery, particularly liver transplants.
- (41) bacteria known to cause ulcers and other stomach related diseases.
- (90) treatments for infectious diseases, particularly those that effect the lungs, the urinary tract, or the pancreas.
- (155) genetic analysis of diseases such as arthritis, epilepsy, and diabetes.
- (68) diagnosis and etiology of alzheimer's disease.
- (30) the use of sentinel node biopsy in the treatment of cancers.
- (42) pain intensity (e.g., postoperative, injury, etc.) and pain relief.
- (51) treatments for orthopedic injuries.
- (87) postoperative complications.

Medicines and Medical Treatment Equipment

- (58) hospital intensive care unit procedures and postoperative recovery, particularly for those recovering from aneurysm surgery.
- (44) coronary artery bypass technologies and procedures.
- (31) treatments for cerebrovascular diseases, such as alzheimer's disease and dementia, with galantanim.
- (36) seizures and ischemic strokes.

- (53) cardiac arrest resuscitation equipment such as defibrillation electrodes, and clinical drugs.
- (46) astronomy, especially asteroid orbits and lunar orbits.
- (52) the cardiovascular system, blood pressure and treatments for these conditions.
- (56) cardiovascular / circulatory system and blood flow.
- (42) clinical trials of various prescription cholesterol drugs, especially gemfibrozil, celiprolol, and pravastatin.
- (42) clinical trials using mifepristone for abortions.
- (85) clinical trials using drugs and placebos (for various health conditions such as asthma, etc.) to ascertain effectiveness.
- (86) clinical trials using medicines (primarily insulin), and placebos for various conditions, but particularly for stabilizing blood glucose levels.

Healthcare, Long Term Medical Conditions, Mental Disorders

- (51) various aspects of physical activity / exercise.
- (33) various aspects of physical training including physical therapy.
- (70) the skeletal system and various fracture healing techniques.
- (48) clinical studies of menopausal women.
- (55) dietary patterns and nutritional contents of foods.
- (44) biological response of fish to drastic dietary change, particularly feed reduction and fasting. Also proteins such as leptin, and hormones that regulate weight and fat concentration for humans.
- (72) the effect of clinical drugs, diet / lifestyle on ldl and hdl cholesterol levels.
- (66) the correlation between blood sugar / glucose levels and obesity.
- (92) diabetes, with particular emphasis in production of insulin and its effect on blood glucose levels.
- (36) causes and risk factors for heart disease.
- (61) mortality risk factors, particularly for risk of stroke or cardiovascular diseases.
- (119) comparing the effects of lifestyles (smoking, alcohol consumption) on the health of men and women.
- (74) socioeconomic factors in Europe and Finland.
- (45) physical and mental health issues, particularly balance and posture issues for the elderly.
- (69) risk factors for lung cancer and asthma, such as smoking, exposure to asbestos, etc.

- (52) classifying / identifying symptoms for medical conditions, particularly temporomandibular disorder (TMD), and asthma.
- (52) organization cultures, climates, and employee health.
- (40) (lifestyle) factors and behaviors contributing to the health of Finnish adolescents, especially smoking (tobacco) and alcohol.
- (90) the health care system in Finland.
- (40) healthcare and health welfare in Finland, with emphasis on nursing and ethics.
- (49) dementia, and available nursing and patient care in Finland.
- (23) assessing and identifying factors (e.g., medical disorders, mental disorders, etc.) that can be correlated with risk for suicide.
- (65) depression, schizophrenia and other mental disorder.
- (28) schizophrenia and other central nervous system disorders.

Pregnancy and Neuropsychological

- (59) various aspects of pregnancy and child birth in Finland.
- (37) human pregnancies and various factors effecting fetal development.
- (30) vaccines for pneumococcal serotype.
- (33) the treatment of an ear infection disease that is prevalent to children.
- (76) sociological aspects of family life.
- (133) health, medicines, etc, for school aged children.
- (39) the development of children's neural systems and their language functions.
- (51) the neural system's underlying language functions.
- (83) the cortex, and methods for visualizing brain activity such as magnetic resonance imaging.
- (51) linguistics / phonological and semantic processes of words.
- (40) ecological species.
- (13) equilibrium thermodynamic descriptions of (copper) ternary material systems.
- (32) sleep behavior and sleep disorders, and measurement of sleep spindle.

SUMMARY AND CONCLUSIONS

The main objective of this study was to assess the technical core competencies of Finland. This was accomplished using partitional clustering and bibliometrics on articles retrieved from the Science Citation Index. At the highest taxonomy level, there appear to be four major research core competencies, based on level of activity:

- Physics/ Information/ Mathematics/ Economics includes about 28% of Finnish research;
- Materials/Chemistry/ Environment/ Ecology includes about 28%;
- Biology/ Biochemistry/ Medicine covers about 17%; and
- Clinical Medicine includes about 27%.

Bibliometrics were performed in detail on four lower level research core competencies identified by analysis of the overall technical taxonomy:

- Wireless Network and Mobile Communications,
- Signal Processing,
- Materials Science and Engineering, and
- Chemistry.

Wireless Network and Mobile Communications

- Journals containing the most papers appear to be applied journals, split between communications and networks.
- Journals addressing more fundamental issues or topics from other disciplines are not represented. IEEE journals are represented most prominently.
- Nokia and Helsinki University of Technology are by far the leaders for Wireless Network and Mobile Communications, with four institutions constituting the second tier: Oulu University, Tempere University of Technology, VTT, University of Helsinki.
- There are nine universities, four research centers, and two industrial organizations. This distribution reflects an applied research/ technology development emphasis and is consistent with the journal section conclusions.
- The major collaborators in this field are the advanced Western countries and Japan.

Signal Processing

- Journals containing the most papers are applied journals, drawing from a number of different applications areas.
- Journals addressing more fundamental issues are not represented.
- Helsinki University of Technology and Tampere University of Technology are the leading institutions for Signal Processing, with two institutions constituting the second tier: University of Oulu and University of Helsinki.
- Four co-authoring countries (USA, England, France, Sweden) predominate in this field, accounting for 44% of the collaborative efforts. The main collaborators are the advanced Western countries.

Materials Science and Engineering

- Articles are published primarily in physics journals, although a small percentage of materials articles appear in materials and chemistry journals.
- Journals addressing more fundamental issues, or topics from other disciplines, are not represented.
- Helsinki University of Technology and University of Helsinki constitute the first tier for Materials Science and Engineering, with four universities and one research institute constituting the second tier (University Turku, VTT, Tampere University Technology, University Oulu, and Abo Akad University).
- The USA is the predominant Materials research partner, with Germany, Sweden, and England constituting the second tier.

Chemistry

- The topical coverage of Chemistry journals is very broad, and the journals are a mix of basic (Journal of Biological Chemistry, Journal of Chemical Physics, Inorganic Chemistry) and applied (Applied Catalysis A, Nordic Pulp and Paper Research Journal, Industrial and Engineering Chemistry Research).
- University of Helsinki is the leading Chemistry institution by far, with four institutions comprising the next tier (University Turku, Helsinki University of Technology, University Kuopio, and University Oulu), and two institutions that constitute the third major tier (Abo Akad University, Tampere University).
- All of the institutions are universities.

- The major collaborating country in Chemistry is the USA, with the next tier having three countries (Sweden, Germany, England).
- The listing of top twenty countries aggregates various groupings (North American developed, Scandanavian countries, regional [Russia, Poland], developed Western democracies, and Finno-Ugric [Hungary, Estonia]).

Integration of these bibliometrics results across the four technical thrusts studied leads to the following general conclusions:

- The journals containing the most papers are mainly applied across the four core competencies selected, with some basic research journal representation in Chemistry.
- University of Helsinki is in the top tier of prolific institutions in the more fundamental research areas, and is in the second tier in the more applied areas.
- Helsinki University of Technology is in the top tier in more applied areas, and in the second tier in more fundamental areas.
- University of Oulu is in the second tier in all four core technology areas selected.
- Nokia is the only company to appear in a first tier (and only in the Wireless Network and Mobile Communications thrust).
- VTT appears in a second tier (in the WNMC and MATLS thrusts).
- The main country collaborators with Finland are the US (all four thrust areas), England (all four thrust areas), Germany (three thrust areas), Sweden (three thrust areas), France (two thrust areas).

The clustering appears useful for generating the structure of a country's S&T, while the bibliometrics appears useful for identifying Centers of Excellence and prolific performers for specific technology areas. Continual upgrades in the clustering algorithms ensure that the accuracy of the clusters and categories will continue to improve.

Two citation-based approaches were used for assessing Finland's technical thrust areas. One approach is based on themes from document clustering and compares the four themes used for the bibliometrics analysis. They are important research areas for Finland. The other approach is based on phrase frequency analysis, and identifies a number of pervasive technical themes that exhibit relatively high activity (numbers of articles published). For

both approaches, the Finnish results are compared with those of two similar countries, Norway and Denmark. All three are small Scandanavian countries with similar populations and GDP. All three are viewed as advanced Western nations.

Document Clustering-based Citations

All the research articles on the four areas selected from the document clustering-based taxonomy published in 1999 with at least one author with a Finland or Norway or Denmark address were retrieved, and their citations examined. Then, global citation statistics were tabulated and compared for each country for each technology.

Wireless Networks and Mobile Communications

The number of Finnish articles in Wireless Networks and Mobile Communications published in both 1999 and 2004 is almost an order of magnitude larger than that from Norway or Denmark. This investment reflects an obvious strategic priority for Finland, and is one characteristic of a true core competency.

Because of the small number of data points for Norway and Denmark for the citation comparison vintage year (1999), the citation comparison statistics for Wireless Networks and Mobile Communications have limited validity. For example, the median of the top twenty percent of articles for Norway and Denmark reflected one article for each country. In addition, many of the Norway and Denmark articles did not focus on development of the wireless/ mobile technology but rather on consequences such as health effects of microwave radiation. Finland did have three highly cited papers (>40), but the relative significance of this number could not be ascertained due to the poor statistics for the other two countries.

To place this result in a larger context, the citation performance of the USA in Wireless Networks and Mobile Communications was examined for articles published in 1999 and compared to that of Finland.

- There were 943 USA articles published in 1999 (compared to Finland's 106).
- The median of the top twenty percent USA article citations was 32 (compared to Finland's 35).
- The median of all the USA article citations was 5 (compared to Finland's 2).

- Finally, the number of USA articles with forty or more citations was 69 (compared to Finland's 3).

Thus, in relative terms, Finland's research output in Wireless Networks and Mobile Communications reflects a higher priority than the USA's, but in absolute terms, the USA is dominant by a large margin.

Signal Processing

In the Signal Processing category, Norway has increased its output substantially since 1999, while Finland's output has increased moderately, and Denmark's has remained the same. Denmark exceeds Finland and Norway noticeably in all the citation metrics used. Most importantly, Denmark has a higher fraction of the most highly cited articles. Thus, while Signal Processing may be a core Finnish research area due to volume of activity, it does not offer a core research impact advantage over Denmark.

Materials

In the Materials category, Finland and Denmark are the clear leaders over Norway in numbers of research articles published, with Finland holding a slight edge. Both Finland and Denmark have increased publication productivity by about a third since 1999, while Norway's increase has been about half that amount. As in the Signal Processing thrust, Denmark outperforms Finland in all the citation metrics used, most importantly in the number of highly cited papers. Again, as in the Signal Processing thrust, Materials research may be a core Finnish research area due to volume of activity, but it certainly does not offer a core research impact advantage over Denmark.

Chemistry

In the Chemistry category, Finland and Denmark are the clear leaders in numbers of publications and are relatively close to each other. Again, Denmark is the clear leader in all the citation metrics used, with moderately less of the commanding lead shown in Materials research.

Three caveats are in order here. The first concerns the relation between citations, research impact, and core competencies. Poor citation performance can reflect:

- Poor intrinsic quality, and/ or
- Low circulation journals, and/ or

- Low research activity in field, and or
- More applied focus, reducing the number of extra-discipline researchers available to cite, and/ or
- Other characteristics.

In particular, a country could produce high quality but very applied research in a technical area. The area's output might receive low citations in aggregate but still be a national core competence. In order to distinguish the reasons for poor citations exhibited by different countries in different categories, one would need to read substantial numbers of research articles produced by the countries in the field of interest and analyze them for intrinsic quality and level of development. That was beyond the scope of the present study.

The second caveat concerns the aggregation level of the technologies examined. The conclusions above about dominance are for the thrust areas defined by the relatively broad queries shown. ***Even though one country may dominate on the thrust area in aggregate (as defined by the total query), another country could conceiveably lead in one of the sub-thrust areas.***

The third caveat concerns the selection of phrases. The selection was based on emphasis in the Finland research output database. A priori, one would expect these to be Finnish research priority areas and would expect Finland's performance relative to other countries to be somewhat higher than average. Any poor relative performance by Finland in these Finnish thrust areas should be cause for concern.

In the above four thrust areas analyzed, Finland has two characteristics of a true core competency: an identifiable cohesive thrust and identifiable critical mass (numbers of publications, relative to Norway and Denmark). However, for the latter three technical thrust areas, it did not have the aggregate citation impact of Denmark. The differences in relative publication and citation attributes across the four technical areas for the three countries show the necessity for comparing countries at the critical technology level rather than at the aggregate national level (King, 2004).

<u>Phrase Frequency-based Citations</u>

A phrase frequency analysis was performed on the contents of the Abstracts of all the Finland records retrieved. The highest frequency technical phrases

were identified. Seventeen of these phrases were selected from the areas of Physical, Environmental, and Life Sciences, with emphasis on phrases different from those used for the core competency queries. No high frequency phrases of adequate specificity could be identified from the Engineering Sciences area. Each phrase (in some cases, combinations of similar phrases) was entered into the SCI search engine, and records were retrieved for Finland, Norway, and Denmark for 1999. Citation indicators were applied to each phrase's retrievals, and the results compared.

The phrase frequency approach shows:

- Finland to be the clear winner in most of the citation metrics in most of the Life Sciences themes,
- Denmark to be the winner in most Environmental Sciences themes,
- Denmark to be a clear winner in four of the seven Physical Sciences themes,
- Finland to be a modest winner in one Physical Sciences theme (Electron Microscopy), and
- Finland and Denmark to be essentially tied in three Physical Sciences themes.

The caveats expressed previously about the interpretation of citations and thematic aggregation apply to these results as well.

The real surprise in these results is the citation performance of Norway. While Norway has the smallest population of the three countries, it has the largest GDP. ***In most cases (but not all), Norway had the lowest research activity, and in almost all cases, had the lowest citation impact.*** It should be re-stated that the phrases selected for the present country comparison analysis, whether from the document clustering-based queries or the phrase frequency-based queries, were based on emphasis in the Finland data base. Whether Norway would perform better if phrases were selected from an analysis of its own database is unknown. As will be shown in the next paragraph, its gross citation numbers do not offer much encouragement, but as this study has emphasized repeatedly, the dis-aggregated critical technologies are of most interest.

Computations were performed to identify the effect of collaborating countries on the research impact of articles with Finnish authors There are

four separate groups of results generated. The first group, Total Country Results, is based on all articles from either 2004 or 1999 that had at least one author with the address of the country of interest (i.e., Finland or Norway or Denmark). The second group, Country Only (No Other Countries), is based on all articles from 1999 with at least one author with the address of the country of interest, excluding all authors with an address from any of Finland's 25 top country collaborators. The purpose of this group is to identify citation characteristics of papers from only the country of interest (Finalnd, Norway, or Denmark).

The third group, Country and USA Only (No Other Countries) is based on all articles from 1999 with at least one author with the address of the country of interest and at least one author with a USA address, excluding all other authors with an address from any of Finland's 25 top country collaborators. The purpose of this group is to identify the impact of adding only USA co-authors to those of the country of interest. The final group, Country and USA, is based on all articles from 1999 with at least one author from the USA and one author from the country of interest. Papers with authors from other countries are not excluded, as they were in the third group.

The summary findings are as follows:

- In the first group, Total Country Results, the gross citation impacts of Finland and Denmark are quite similar, and substantially larger than that of Norway.
- In the second group, Country Only, Denmark has modestly better citation performance than Finland, which in turn has modestly better citation performance than Norway.
- In the third group, Country and USA Only (No Other Countries), Finland and the USA outperform Denmark and the USA. Both groups substantially outperform Norway and the USA.
- In the fourth group, Country and USA, multi-country papers including Finland and USA very slightly outperform multi-country papers including Denmark and USA, which in turn moderately outperform multi-country papers including Norway and the USA.
- Adding the USA only to Finland/ Norway/ Denmark papers increases the citation performance substantially, and adding more countries enhances performance modestly for Finland, somewhat more for Denmark, and substantially for Norway.

The citation-based approaches appear very useful for comparing research impact among countries, but substantial reading of research outputs is required for proper interpretation of citation results.

REFERENCES

Bostian CW, Brandon WT, Mac Rae AU, Mahle CE, Townes SA . Key technology trends - Satellite systems. Space Communications. 16 (2-3): 97-124 2000.

Campbell, R., H.D. Balzer, J. Berliner, R. Dobson, and P. Gregory. "Soviet Science and Technology," Foreign Applied Sciences Assessment Center, October 15, 1985.

Cohn, M., L.W. Craver, A. Gersho, T. Lookabaugh, F. Pollara, and M. Vetterli, *Non-US Data Compression and Coding Research*, Gray, R.M., ed November 1993. A Foreign Applied Sciences Assessment Center (FASAC) report prepared for Science Applications International Corporation (SAIC) under U.S. Government sponsorship.

Cutting DR, Karger DR, Pedersen JO, Tukey JW. Scatter/Gather: A cluster-based approach to browsing large document collections. In *Proceedings of the 15th International ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR'92). 1992. 318-329.

Davidson, R.C. M.A. Abdou, L.A. Berry, C.W. Horton, J.F. Lyon, and P.H. Rutherford, Japanese Magnetic Confinement Fusion Research, Foreign Applied Sciences Assessment Center Technical Assessment Report, Science Applications International Corporation, 1990.

Duncan, L. M., F.T. Djuth, J.A. Fejer, N.C. Gerson, T. Hagfors, D.B. Newman, Jr., R.L. Showen, "Soviet Ionospheric Modification Research," Foreign Applied Sciences Assessment Center, Technical Assessment Report 4040, 1988.

Finnish Academy, "Academy selects 23 new centres of excellence in research", Press release 7th June 2005, available at http://www.aka.fi.

Rodan S, Galunic C. More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness. Strategic Management Journal 25 (6): 541-562. 2004

Garfield E. History of citation indexes for chemistry - a brief review. JCICS. 1985; 25(3): 170-174.

REFERENCES

Guha S, Rastogi R, Shim K. CURE: An efficient clustering algorithm for large databases. In *Proceedings of the ACM-SIGMOD 1998 International Conference on Management of Data* (SIGMOD'98). 1998. 73-84.

Prahalad CK, Hamel G. The core competence of the corporation. Harvard Business Review 68 (3): 79-91. May-Jun 1990.

Hutubessy RCW, Hanvoravongchai P, Edejer TTT. Diffusion and utilization of magnetic resonance imaging in Asia. International Journal of Technology Assessment in Health Care. 18 (3): 690-704 SUM 2002.

Karypis G, Han EH, Kumar V. Chameleon: A hierarchical clustering algorithm using dynamic modeling. IEEE Computer: Special Issue on Data Analysis and Mining. 1999. 32(8). 68--75.

Karypis G. CLUTO—A clustering toolkit. http://www.cs.umn.edu/~cluto. 2004.

King, D. A. The scientific impact of nations. *Nature* 430 (2004) 311-316.

Klinger, A., editor, Klinger, A., et. al., "Soviet Image Pattern Recognition Research," Jan. 1990, Foreign Applied Sciences Assessment Center, *Science Applications International Corp.*, 10260 Campus Point Drive, San Diego, CA 92121, and 1710 Goodridge Drive, McLean VA 22102.

Kostoff, R. N., "Database Tomography for Technical Intelligence: Comparative Analysis of the Research Impact Assessment Literature and the Journal of the American Chemical Society:, Scientometrics, 40:1, 1997.

Kostoff, R. N., Eberhart, H. J., and Toothman, D. R. "Database Tomography for Technical Intelligence: A Roadmap of the Near-Earth Space Science and Technology Literature". Information Processing and Management. 34:1. 1998a.

Kostoff RN. The use and misuse of citation analysis in research evaluation. Scientometrics 1998b; 43:1: 27-43.

Kostoff, R. N., Eberhart, H. J., and Toothman, D. R. "Hypersonic and Supersonic Flow Roadmaps Using Bibliometrics and Database

Tomography". Journal of the American Society for Information Science. 50:5. 427-447. 15 April 1999.

Kostoff, R. N., Braun, T., Schubert, A., Toothman, D. R., and Humenik, J. "Fullerene Roadmaps Using Bibliometrics and Database Tomography". Journal of Chemical Information and Computer Science. 40:1. 19-39. Jan-Feb 2000a.

Kostoff, R. N., Green, K. A., Toothman, D. R., and Humenik, J. "Database Tomography Applied to an Aircraft Science and Technology Investment Strategy". Journal of Aircraft, 37:4. 727-730. July-August 2000b.

Kostoff, R. N., and DeMarco, R. A. "Science and Technology Text Mining". Analytical Chemistry. 73:13. 370-378A. 1 July 2001a.

Kostoff, R. N., Del Rio, J. A., García, E. O., Ramírez, A. M., and Humenik, J. A. "Citation Mining: Integrating Text Mining and Bibliometrics for Research User Profiling". JASIST. 52:13. 1148-1156. 52:13. November 2001b.

Kostoff, R. N., Tshiteya, R., Pfeil, K. M., and Humenik, J. A. "Electrochemical Power Source Roadmaps using Bibliometrics and Database Tomography". Journal of Power Sources. 110:1. 163-176. 2002.

Kostoff, R. N. "Text Mining for Global Technology Watch". In Encyclopedia of Library and Information Science, Second Edition. Drake, M., Ed. Marcel Dekker, Inc. New York, NY. 2003. Vol. 4. 2789-2799.

Kostoff, R. N., Shlesinger, M., and Tshiteya, R. "Nonlinear Dynamics Roadmaps using Bibliometrics and Database Tomography". International Journal of Bifurcation and Chaos. 14:1. 61-92. January 2004a.

Kostoff, R. N., Shlesinger, M., and Malpohl, G. "Fractals Roadmaps using Bibliometrics and Database Tomography". Fractals. 12:1. 1-16. March 2004b.

Kostoff, R.N., Bedford, C.W., Del Rio, J. A ., Cortes, H., and Karypis, G. "Macromolecule Mass Spectrometry: Citation Mining of User Documents".

Journal of the American Society for Mass Spectrometry. 15:3. 281-287. March 2004c.

Kostoff, R. N., Buchtel, H., Andrews, J., and Pfeil, K. "The hidden structure of neuropsychology: Text Mining of the Journal *Cortex*: 1991-2001". Cortex. 41:2. 103-115. April 2005a.

Kostoff, R. N., Del Rio, J. A., Smith, C., Smith, A., Wagner, C.S., Malpohl, G., Karypis, G., and Tshiteya, R. "The Structure and Infrastructure of Mexico's Science and Technology". Technological Forecasting and Social Change. 72:7. August 2005b.

Lanzerotti, L. J., R. C. Henry, H. P. Klein, H. Masursky, G. A. Paulikas, F. L. Scarf, G. A. Soffen, and Y. Terzian, "Soviet Space Science Research," FASAC Technical Assessment Report FASAC-TAR-3060, Foreign Applied Sciences Assessment Center, 1986.

Leneman B . Automation In Soviet Industry, 1970-1983 - An Assessment Of The Present State Of Robot-Technology. Revue D Etudes Comparatives Est-Ouest. 15 (1): 75-112 1984.

MacRoberts M, MacRoberts B. Problems of citation analysis. Scientometrics 1996. 36(3): 435-444.

McIntire LV . WTEC panel report on tissue engineering (Reprinted). Tissue Engineering. 9 (1): 3-7 FEB 2003.

Ministry of Education, "Opetusministeriön toiminta- ja taloussuunnitelma 2003-2006", 2003, pp. 49, available at http://www.minedu.fi/julkaisut/toim/tts2003-2006.html

Mooney B, Seymour R . WTEC panels survey Russian maritime technologies. Marine Technology Society Journal. 30 (1): 71-72 SPR 1996.

Rantanen J, "Yliopistojen ja ammattikorkeakoulujen tutkimuksen rakenneselvitys," Opetusministeriön työryhmämuistioita ja selvityksiä, 2004:36, Opetusministeriö, p. 204, available at http://www.minedu.fi/julkaisut/index.html.

REFERENCES

Rasmussen E. Clustering Algorithms. In W. B. Frakes and R. Baeza-Yates (eds.). Information Retrieval Data Structures and Algorithms. 1992. Prentice Hall, N. J.

Spencer, W.J., J.Y. Chen, A. Chiang, W. Frieman, E.S. Kuh, J.L. Moll, R.F. Pease, and K.C. Saraswat, "Chinese Microelectronics," Foreign Applied Sciences Assessment Center Technical Assessment Report, Science Applications International Corporation, April 1989.

Stares P . United-States and Soviet Military Space Programs - A Comparative-Assessment. Daedalus. 114 (2): 127-145 1985.

Statistics Finland, ”R&D personnel (head count) by sector in 1991-2003”, Statistics Finland, available at http://www.stat.fi/tk/yr/tttiede_rd3_en.html.

Steinbach M, Karypis G, Kumar V. A comparison of document clustering techniques. Technical Report #00--034. 2000. Department of Computer Science and Engineering. University of Minnesota.

Willet P. Recent trends in hierarchical document clustering: A critical review. Information Processing and Management. 1988. 24:577-597.

Zamir O, Etzioni O. Web document clustering: A feasibility demonstration. In: Proceedings of the 19th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'98). 1998. 46-54.

Zhao Y, Karypis G. Empirical and theoretical comparisons of selected criterion functions for document clustering. Machine Learning 55 (3): 311-331. 2004.

APPENDIX 1 – RESEARCH BIBLIOMETRICS

1-A. Signal Processing

Based on the computational linguistics (clustering) results, Signal Processing is an important thrust area of Finnish research. Starting with the words generated by the clustering algorithm for the Signal Processing cluster, an iterative feedback approach was used to generate the following comprehensive query for this research in Finland:

"SIGNAL PROCESSING OR DATA FORMAT CONVERTER* OR BANDPASS SAMPL* OR ((SIGNAL OR DIGITAL) AND (FILTER* OR CODE OR CODING OR TRANSMIT* OR AUDIO OR VIDEO OR SPEECH OR IMAGE OR COMMUNICATION OR GEOPHYSIC* OR ACOUSTIC* OR SONAR OR RADAR OR MUSIC* OR ANALOG CONVERTER* OR WAVEGUIDE OR WAVE GUIDE OR RECEIVER*))"

The query was inserted into the Science Citation Index, and about 170 records were recovered for the period 2003-2004 (approximately one percent of the 15700 SCI records in Signal Processing for this period). The bibliometrics analysis was performed on these records.

1-A-1. Prolific Authors

Table 1-A-1 lists the thirteen most prolific authors in Finnish Signal Processing research, including their institutions. Two institutions predominate: Helsinki University of Technology and Tampere University of Technology. Twelve are universities, and one company (foreign) is listed.

TABLE 1-A-1 – MOST PROLIFIC FINNISH SIGNAL PROCESSING AUTHORS

AUTHOR	INSTITUTION	#PAPERS
VÄLIMÄKI--V	HELSINKI UNIVERSITY TECHNOLOGY	11
KARJALAINEN--M	HELSINKI UNIVERSITY TECHNOLOGY	6
HALONEN--KAI	HELSINKI UNIVERSITY TECHNOLOGY	5
ALKU--P	HELSINKI UNIVERSITY TECHNOLOGY	4
RENFORS--M	TAMPERE UNIVERSITY TECHNOLOGY	4
ABE—T	NTT DOCOMO (JAPAN)	3
BACKSTROM--T	HELSINKI UNIVERSITY TECHNOLOGY	3

DUMITRESCU--B	TAMPERE UNIVERSITY TECHNOLOGY	3
ERKUT--C	HELSINKI UNIVERSITY TECHNOLOGY	3
ESQUEF--PAA	HELSINKI UNIVERSITY TECHNOLOGY	3
LAMMINEN--H	TAMPERE UNIVERSITY TECHNOLOGY	3
MATSUMOTO--T	UNIVERSITY OULU	3
PARKKINEN--J	UNIVERSITY JOENSUU	3

1-A-2. Prolific Journals

Table 1-A-2 lists the twenty most prolific journals containing Finnish Signal Processing research papers. These appeared to be applied journals, drawing from a number of different applications areas. Journals addressing more fundamental issues are not represented.

TABLE 1-A-2 – MOST PROLIFIC JOURNALS – FINNISH SIGNAL PROCESSING RESEARCH

JOURNAL	#PAPERS
EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING	10
JOURNAL OF THE AUDIO ENGINEERING SOCIETY	6
IEEE SIGNAL PROCESSING LETTERS	6
IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I-REGULAR PAPERS	3
IMAGE ANALYSIS, PROCEEDINGS	3
ANNALES GEOPHYSICAE	3
IEEE JOURNAL OF SOLID-STATE CIRCUITS	3
IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING	3
SIGNAL PROCESSING	3
SCANDINAVIAN JOURNAL OF FOREST RESEARCH	3
IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II-ANALOG AND DIGITAL SIGNAL PROCESSING	3
ELECTRONICS LETTERS	3
MARINE BIOLOGY	2
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SECTION A-ACCELERATORS SPECTROMETERS DETECTORS AND	2
EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS	2
IEEE TRANSACTIONS ON SIGNAL PROCESSING	2
IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY	2
INVESTIGATIVE OPHTHALMOLOGY & VISUAL SCIENCE	2
COMPUTER SYSTEMS: ARCHITECTURES, MODELING, AND SIMULATION	2
IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS	2
IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING	2
NEUROIMAGE	2

1-A-3. Prolific Institutions and Countries

This section identifies the most prolific institutions producing Finnish-authored Signal Processing papers, and the countries of the most prolific collaborators with Finnish authors of Signal Processing papers.

Table 1-A-3A contains a list of the twelve most prolific institutions.for Finnish-authored Signal Processing papers, and Table 1-A-3B contains a list of the fifteen most prolific countries associated with Finnish-authored Signal Processing papers. Helsinki University of Technology and Tampere University of Technology are the leaders, with two institutions constituting the second tier: University of Oulu and University of Helsinki. Four co-authoring countries predominate (USA, England, France, Sweden), accounting for 44% of the collaborative efforts. The main collaborators are the advanced Western countries.

TABLE 1-A-3A MOST PROLIFIC INSTITUTIONS–FINNISH SIGNAL PROCESSING RESEARCH

INSTITUTION	#PAPERS
HELSINKI UNIV TECHNOL	46
TAMPERE UNIV TECHNOL	39
UNIV OULU	28
UNIV HELSINKI	26
UNIV JOENSUU	8
VTT TECHNICAL RESEARCH CTR OF FINLAND	8
UNIV TURKU	7
UNIV JYVASKYLA	6
NOKIA RES CTR	5
FINNISH FOREST RES INST	4
LAPPEENRANTA UNIV TECHNOL	4
UNIV KUOPIO	3

Formatted: All caps

TABLE 1-A-3B MOST PROLIFIC COUNTRIES–FINNISH SIGNAL PROCESSING RESEARCH

COUNTRY	#PAPERS
FINLAND	171
USA	12
ENGLAND	8
FRANCE	6
SWEDEN	6
CANADA	4

ITALY	4
JAPAN	4
DENMARK	3
BRAZIL	2
GERMANY	2
POLAND	2
RUSSIA	2
WALES	2
YUGOSLAVIA	2

1-B. Materials Science and Engineering

Based on the computational linguistics (clustering) results, Materials Science and Engineering (MSE) is an important thrust area of Finnish research. Starting with the words generated by the clustering algorithm for the MSE cluster, an iterative feedback approach was used to generate the following comprehensive query for this research in Finland:

(NANO* NOT (NANOSECOND* OR NANOLITER* OR NANOMOLAR* OR NANOBACTERIA OR NANO3 OR NANOFILTRATION OR NANOCOLLOID*)) OR SMART MATERIAL* OR METALS OR METALLIC OR ALLOY* OR CORROSI* OR BIOCHEMICAL MATERIAL* OR THIN FILM* OR CERAMIC* OR (MATERIAL* SAME (COPOLYMER* OR POLYMER* OR FILM*)) OR (COMPOSITE* SAME (REINFORCE* OR BOND* OR RESIN* OR STRENGTH OR FLEXURE OR GLASS OR FRACTURE OR FIBER OR SUBSTRATE OR PROPERTIES OR MATERIAL*)) OR ATOMIC FORCE MICROSCOPY OR SCANNING ELECTRON MICROSCOPY OR SCANNING TUNNELING MICROSCOPY OR CHEMICAL VAPOR DEPOSITION OR STEEL OR STEELS OR SLAG OR TENSILE TEST* OR X-RAY DIFFRACTION OR ATOMIC LAYER OR SURFACE ROUGHNESS

The query was inserted into the Science Citation Index, and 1091 records were recovered for the period 2003-2004. The bibliometrics analysis was performed on these records.

1-B-1. Prolific Authors

Table 1-B-1 lists the twenty most prolific authors in Finnish MSE research, including their institutions. Two institutions predominate: Helsinki University of Technology and University Helsinki.

TABLE 1-B-1 MOST PROLIFIC FINNISH MSE AUTHORS

AUTHOR	INSTITUTION	#PAPERS
LESKELA--M	UNIVERSITY HELSINKI	37
RITALA--M	UNIVERSITY HELSINKI	33
VALLITTU--PK	UNIV TURKU	33
NORDLUND--K	UNIVERSITY HELSINKI	29
KEINONEN--J	UNIVERSITY HELSINKI	27
NIEMINEN--RM	HELSINKI UNIV TECHNOLOGY	23
LINDROOS--VK	HELSINKI UNIV TECHNOLOGY	22
ULLAKKO--K	HELSINKI UNIV TECHNOLOGY	21
NIINISTO--L	HELSINKI UNIV TECHNOLOGY	20
KUKLI--K	UNIVERSITY TARTU	18
SAJAVAARA--T	UNIVERSITY HELSINKI	17
SODERBERG--O	HELSINKI UNIV TECHNOLOGY	17
HECZKO--O	HELSINKI UNIV TECHNOLOGY	16
LINDEN--M	ABO AKAD UNIVERSITY	16
PESSA--M	TAMPERE UNIVERSITY TECHNOLOGY	16
TORMALA--P	TAMPERE UNIVERSITY TECHNOLOGY	16
LASSILA--LVJ	UNIVERSITY TURKU	15
MURZIN--DY	ABO AKAD UNIVERSITY	15
SALMI--T	ABO AKAD UNIVERSITY	14
SERIMAA--R	UNIVERSITY HELSINKI	14

1-B-2. Prolific Journals

Table 1-B-2 lists the seventeen most prolific journals containing Finnish MSE research papers. These appear to be mainly physics journals, with some materials and chemistry journals at lower frequencies. Journals addressing more fundamental issues, or topics from other disciplines, are not represented.

TABLE 1-B-2 MOST PROLIFIC JOURNALS – FINNISH MSE RESEARCH

JOURNAL	#PAPERS
PHYSICAL REVIEW B	37
JOURNAL OF APPLIED PHYSICS	24
JOURNAL DE PHYSIQUE IV	22
APPLIED PHYSICS LETTERS	20

THIN SOLID FILMS	17
LANGMUIR	16
SURFACE SCIENCE	14
NUCLEAR INSTRUMENTS & METHODS IN PHYSICS RESEARCH SECTION B-BEAM INTERACTIONS WITH MATERIALS AND ATO	14
JOURNAL OF PHYSICS-CONDENSED MATTER	12
APPLIED SURFACE SCIENCE	12
MATERIALS SCIENCE AND ENGINEERING A-STRUCTURAL MATERIALS PROPERTIES MICROSTRUCTURE AND PROCESSING	11
JOURNAL OF PHYSICAL CHEMISTRY B	11
JOURNAL OF CRYSTAL GROWTH	10
JOURNAL OF THE ELECTROCHEMICAL SOCIETY	10
CHEMICAL VAPOR DEPOSITION	9
APPLIED CATALYSIS A-GENERAL	9
JOURNAL OF MATERIALS SCIENCE-MATERIALS IN ELECTRONICS	9

1-B-3. Prolific Institutions and Countries

This section identifies the most prolific institutions producing Finnish-authored MSE papers, and the countries of the most prolific collaborators with Finnish authors of MSE papers.

Table 1-B-3A contains a list of the twenty most prolific institutions.for Finnish-authored MSE papers, and Table 1-B-3B contains a list of the twenty most prolific countries associated with Finnish-authored MSE papers Helsinki University of Technology and University of Helsinki constitute the first tier, with four universities and one company constituting the second tier (University Turku, VTT, Tampere University Technology, University Oulu, Abo Akad University). The USA is the predominant research partner, with Germany, Sweden, and England constituting the second tier.

TABLE 1-B-3A MOST PROLIFIC INSTITUTIONS–FINNISH MSE RESEARCH

INSTITUTION	#PAPERS
HELSINKI UNIV TECHNOL	273
UNIV HELSINKI	213
UNIV TURKU	124
VTT	117
TAMPERE UNIV TECHNOL	102
UNIV OULU	93
ABO AKAD UNIV	83
UNIV JOENSUU	44
UNIV JYVASKYLA	43
UNIV KUOPIO	31

LAPPEENRANTA UNIV TECHNOL	19
UNIV TARTU	18
TAMPERE UNIVERSITY	16
FINNISH FOREST RES INST	15
UNIV GRONINGEN	14
FINNISH METEOROL INST	13
RUSSIAN ACAD SCI	12
GEOL SURVEY FINLAND	12
POLISH ACAD SCI	11
UNIV CALIF SANTA BARBARA	10

TABLE 1-B-3B MOST PROLIFIC COUNTRIES – FINNISH MSE RESEARCH

COUNTRY	#PAPERS
FINLAND	1091
USA	90
GERMANY	79
SWEDEN	52
ENGLAND	50
NETHERLANDS	39
FRANCE	38
RUSSIA	36
ESTONIA	27
POLAND	27
SPAIN	26
JAPAN	19
AUSTRIA	17
BELGIUM	17
HUNGARY	16
CZECH REPUBLIC	14
ITALY	13
DENMARK	12
IRELAND	11
PEOPLES R CHINA	11

1-C. Chemistry

Based on the computational linguistics (clustering) results, Chemistry is an important thrust area of Finnish research. Starting with the words generated by the clustering algorithm for the Chemistry cluster, an iterative feedback approach was used to generate the following comprehensive query for this research in Finland:

"CHEMISTRY OR BIOCHEMISTRY OR CHEMICAL OR BIOCHEMICAL OR CATALYST* OR CATALYTIC OR CATALYSIS OR (REACTION* AND (ENANTIOSELECT* OR KINETIC* OR HYDROGEN OR OXYGEN OR ACID OR ACIDS OR SOLVENT* OR SOLUTION* OR OXIDATION OR COMPLEX*)) OR (COMPOUND* AND (SYNTHESIS OR DERIVATIVE* OR ACID* OR BOND* OR MOLECUL* OR CLEAVAGE*)) OR (NMR AND (METHYL OR RING* OR COMPOUND* OR STRUCTURE* OR PHENYL OR REACTION* OR ISOMER*)) OR (COMPLEXES AND (LIGAND* OR BOND* OR ION* OR CATION* OR ATOM* OR STRUCTURE* OR ANION*)) OR ((POLYMER* OR CO-POLYMER*) AND (CHAIN* OR BLOCK* OR SYNTHESIS OR COMPLEX*)) OR (SOLUTION* AND (ACID* OR SODIUM OR AQUEOUS OR ION EXCHANGE)) OR (SULFURIC ACID AND WATER) OR LIQUID CHROMATOGRAPH* OR MASS SPECTROMET* OR POLYCYCLIC AROMATIC HYDROCARBON* OR RADICAL SCAVENG*"

The query was inserted into the Science Citation Index, and over 2500 records were recovered for the period 2003-2004. The bibliometrics analysis was performed on these records.

1-C-1. Prolific Authors

Table 1-C-1 lists the thirteen most prolific authors in Finnish Chemistry research, including their institutions. Three institutions predominate: Abo Akad University, Helsinki University of Technology, and University Jyvaskyla.

TABLE 1-C-1 MOST PROLIFIC FINNISH CHEMISTRY AUTHORS

AUTHOR	INSTITUTION	THEME	#PAPERS
MURZIN—DY	ABO AKAD UNIV	CATALYSIS	57
SALMI—T	ABO AKAD UNIV	CATALYSIS, KINETICS	50
LESKELA—M	UNIV HELSINKI	THIN FILMS	47
PIHLAJA—K	UNIV TURKU	ANALYTICAL CHEMISTRY	39
RIEKKOLA--ML	UNIV HELSINKI	SEPARATION	34
KULMALA—M	UNIV HELSINKI	ATMOSPHERIC CHEMISTRY	33
HAUKKA—M	UNIV JOENSUU	SYNTHESIS	31
MAKI-ARVELA--P	ABO AKAD UNIV	CATALYSIS	30
RISSANEN--K	UNIV JYVASKYLA	SYNTHESIS	29
PAKKANEN--TA	FINNISH METEOROL INST	AEROSOLS, NANOSTRUCTURES	28

KUMAR--N	ABO AKAD UNIV	CATALYSIS	24
SILLANPAA--R	UNIV JYVASKYLA	ORGANOMETALLICS	24
VUORINEN--T	UNIV KUOPIO	ENVIRONMENTAL CHEMISTRY	24

1-C-2. Prolific Journals

Table 1-C-2 lists the twenty most prolific journals containing Finnish Chemistry research papers. The topical coverage is very broad, and the journals are a mix of basic (Journal of Biological Chemistry, Journal of Chemical Physics, Inorganic Chemistry) and applied (Applied Catalysis A, Nordic Pulp and Paper Research Journal, Industrial and Engineering Chemistry Research).

TABLE 1-C-2 – MOST PROLIFIC JOURNALS – FINNISH CHEMISTRY RESEARCH

JOURNAL	#PAPERS
JOURNAL OF BIOLOGICAL CHEMISTRY	40
JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY	26
JOURNAL OF CHROMATOGRAPHY A	23
JOURNAL OF CHEMICAL PHYSICS	21
INORGANIC CHEMISTRY	20
FEBS LETTERS	19
APPLIED CATALYSIS A-GENERAL	19
ATMOSPHERIC ENVIRONMENT	19
NORDIC PULP & PAPER RESEARCH JOURNAL	18
HOLZFORSCHUNG	18
INDUSTRIAL & ENGINEERING CHEMISTRY RESEARCH	17
PHYSICAL REVIEW B	16
JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES	15
PHYSICAL CHEMISTRY CHEMICAL PHYSICS	15
ORGANOMETALLICS	15
LANGMUIR	15
CLINICAL CHEMISTRY	15
MACROMOLECULES	14
JOURNAL OF CATALYSIS	14
BOREAL ENVIRONMENT RESEARCH	13

1-C-3. Prolific Institutions and Countries

This section identifies the most prolific institutions producing Finnish-authored Chemistry papers, and the countries of the most prolific collaborators with Finnish authors of Chemistry papers.

Table 1-C-3A contains a list of the nine most prolific institutions.for Finnish-authored Chemistry papers, and Table 1-C-3B contains a list of the eight most prolific countries associated with Finnish-authored Chemistry papers. University of Helsinki is the the leader by far, with four institutions that constitute the next tier (University Turku, Helsinki University of Technology, University Kuopio, and University Oulu), and two institutions that constitute the third major tier (Abo Akad University, Tampere University). All of the institutions are universities.

The major collaborating country is the USA, with the next tier having three countries (Sweden, Germany, England). The top twenty country listing aggregates various groupings (North American developed, Scandanavian countries, regional [Russia, Poland], developed Western democracies, and Finno-Ugric [Hungary, Estonia]).

TABLE 1-C-3A – MOST PROLIFIC INSTITUTIONS – FINNISH CHEMISTRY RESEARCH

INSTITUTION	#PAPERS
UNIV HELSINKI	886
UNIV TURKU	355
HELSINKI UNIV TECHNOL	275
UNIV KUOPIO	257
UNIV OULU	240
ABO AKAD UNIV	220
TAMPERE UNIV	190
UNIV JYVASKYLA	166
UNIV JOENSUU	150

TABLE 1-C-3B – MOST PROLIFIC COUNTRIES – FINNISH CHEMISTRY RESEARCH

COUNTRY	#PAPERS
FINLAND	2534
USA	209
GERMANY	178
SWEDEN	168
ENGLAND	137
RUSSIA	99
NETHERLANDS	97
FRANCE	91
SPAIN	67
ITALY	64
CANADA	57

DENMARK	55
POLAND	52
NORWAY	46
ESTONIA	44
JAPAN	41
BELGIUM	39
HUNGARY	37
SWITZERLAND	36
CZECH REPUBLIC	29

APPENDIX 2 – CITATION COMPARISONS

This appendix will present an approach for comparing two of the technology thrusts common to the Mexico and Finland studies (Mexico was selected for comparison initially because it was the first country assessed by our text mining process). The technical areas selected for demonstrating the approach are thin film research and gene research. They are important research areas globally, and were identifiable thrusts for both Mexico and Finland. However, thin films were higher on Mexico's taxonomy than Finland's, suggesting they are more a core competency for Mexico than Finland. The method used for comparison was to compare the citation results for papers of each of the two technologies published in a vintage year, focusing on the most and least cited papers for each country. This method has been used in other contexts recently, and provides good insights into characteristics of highly and lowly cited papers. The year 1999 was selected as the comparison year. It is of sufficient vintage to allow citations to accumulate, yet sufficiently near-term to insure relevancy. Switzerland was used for normalization purposes, as an example of a high tech country.

All the research articles on thin films and genes published in 1999 with at least one author with a Finland or Mexico (or Switzerland) address were retrieved, and their citations examined. Then, global citation statistics for each country for each technology were tabulated and compared. Finally, the most and least cited articles from each country for each technology were grouped, and a number of attributes compared.

The global comparison component was selected for the following reasons. In evaluating research impact, there are three main criteria to consider: 'right job', 'job right', 'productivity/ progress'. 'Right job' refers to proper selection of the broadest objectives; i.e., is the right study being pursued? "Job right' refers to selection of the best approaches to solving the problem to reach the desired goal. 'Productivity/ progress' refer to whether anything tangible is being accomplished.

A detailed determination of 'right job' using citation statistics would require clustering the vintage papers thematically, examining citation ranges for each cluster (theme), then assuming that those themes that had the highest citations were the 'hot' research areas. The papers that were in the 'hot' clusters would get high ratings for the 'right job' criterion. The 'job right' rating for any of the papers would be determined by its citation position

within any of the clusters. However, for this initial country application of the global comparison, the first two criteria are combined, and the overall citation statistics for all the thin film and gene papers will be compared for the two countries (and Switzerland).

Then, the least and most cited papers from each country will be compared intra-country and inter-country. The attributes compared will be both quantitative and qualitative, to provide a broader perspective on important comparative attributes.

TABLE A2-A – FILMS – CITATION COMPARISON

	FINLAND		**MEXICO**		**SWITZERLAND**	
	MOST CITED	*LEAST* CITED	*MOST* CITED	*LEAST* CITED	*MOST* CITED	*LEAST* CITED
#CITATIONS						
AVERAGE	31.2	0	37	0	87.8	0
MEDIAN	28	0	27.5	0	76.5	0
STUDY TYPE						
EXPERIMENT	9	16	9	15	9	19
THEORY	1	1	1	2	1	1
COMP MODEL	0	2	0	2	0	0
DATA ANALYSIS						
SYSTEMS ANALYSIS	0	1	0	1	0	0
TECHNOLOGY LEVEL						
BASIC RES	7	1	6	2	2	0
APPLIED RES	3	15	4	15	8	16
TECH DEV	0	5	0	3	0	4
TECHNICAL THEME						
ENERGY/ELECTRONIC MATERIALS	0	13	2	9	6	12
THERMAL/DIFFUSION CONTROL	0	2	0	0	0	0
SENSORS AND DETECTORS	1	1	0	0	0	1
MEDICAL	1	0	0	0	2	1
ADHESION/WEAR/MECH. PROP'S	2	1	0	1	0	3
ELECTROCHEM AND CORROSION	1	1	0	3	0	0
GENERAL FILM SYNTH/NANOTECH	5	1	8	3	1	3
OTHER	0	1	0	4	1	0

TABLE A2-B – GENES – CITATION COMPARISON

	FINLAND		**MEXICO**		**SWITZERLAND**	
	MOST CITED	*LEAST* CITED	*MOST* CITED	*LEAST* CITED	*MOST* CITED	*LEAST* CITED
#CITATIONS						
AVERAGE	227.5	0.5	77.4	0	326.9	0
MEDIAN	187.5	0.5	67	0	268	0
STUDY TYPE						

EXPERIMENT	8	14	9	14	10	9
THEORY	0	2	0	1	0	1
COMP MODEL	1	0	0	0	0	0
DATA ANALYSIS (EPIDEM)	1	4	0	5	0	9
SYSTEMS ANALYSIS	0	0	1	0	0	1
TECHNOLOGY LEVEL						
BASIC RES	3	1	4	4	6	3
APPLIED RES	6	18	6	15	4	11
TECH DEV	1	1	0	1	0	6
TECHNICAL THEME						
GENE PHYSIOLOGY	3	0	3	2	4	2
GENE IDENTIFICATION/MAPPING	1	2	1	1	1	1
EPIDEMIOLOGY	6	10	6	11	4	13
FETAL CELLS/DEVELOPMENT	0	2	0	0	0	1
PLANT GENETICS/IMMUNOLOGY:	0	2	0	2	1	0
ANIMAL GENETICS/DISEASES:	0	1	0	4	0	1
OTHER	0	3	0	0	0	2

Table 4A contains the citation comparison results for Films, and Table 4B contains the citation comparison results for Genes.

Films

The categories in Table A2-A will be addressed from the top, proceeding downward. For Citations, Finland and Mexico have about the same median citations for Films, while Switzerland has more than 2.5 times as many median citations. Mexico has about twenty percent more average citations than Finland, showing the presence of more of the very highly cited papers.

For Study Type, most of the papers can be classified as experimental, and for all three countries, ninety percent of the most cited papers were in the experimental category. The numbers in the other categories were too small to draw definitive conclusions, although all papers in the computational modeling and systems analysis categories were in the least cited category. This is in line with other technologies for which citation comparisons have been made.

For Technology Level, most of the highly cited papers for Finland and Mexico were in the basic research category, whereas for Switzerland, most of the highly cited papers were in the applied research category. For Finland and Mexico, there were twice as many least cited papers in the applied research category as most cited, while for Switzerland, the ratio was 50/ 50.

For all three countries, none of the technology development papers were in the highly cited category.

For Technical Theme, most of the effort of all three countries is concentrated in the energy/ electronic materials category, with secondary emphasis on the general film synthesis/ nanotechnology category. In energy/ electronic materials, most of Finland's and Mexico's articles are in the least cited category, whereas Switzerland's papers in this category are divided evenly. In general film synthesis/ nanotechnology, most of Finland's and Mexico's articles are in the most cited category, whereas Switzerland's papers in this category are again about evenly divided.

Thus, Finland and Mexico are quite similar across the categories selected for Films, both are well behind Switzerland in citations received by most highly cited papers, and both seem to have more papers in basic research than Switzerland.

Genes

The categories in Table A2-B will be addressed from the top, proceeding downward. For Citations, Finland has about three times the median of citations of most cited papers as Mexico, and about two thirds the number of Switzerland.

For Study Type, most of the papers can be classified as experimental, and for all three countries, over eighty percent of the most cited papers were in the experimental category. For Finland and Mexico, slightly over half the papers in the experimental category were in the most cited category, whereas for Switzerland, about two thirds of the experimental papers were in the most cited category. The only other category of significance was the data analysis/ epidemiology category, and almost all the papers in this category were in the least cited category. This is in line with other medical disciplines for which citation comparisons have been made.

For Technology Level, most of the papers for Finland and Mexico were in the applied research category, whereas for Switzerland, half the papers were in the applied research category. For all three countries, a moderately higher percentage of the applied research papers were in the least cited category. For all three countries, papers in the basic research categories were

predominantly in the most cited category. Essentially none of the technology development papers were in the highly cited category.

For Technical Theme, most of the effort of all three countries is concentrated in the epidemiology category, with secondary emphasis on the gene physiology category. There is about an even split in the most cited/ least cited categories, For gene physiology, most of the records are in the most cited categories

Thus, Finland and Mexico are quite similar in attributes, with the exception that Finland has substantially more citations, on average, than Mexico.

In summary, for the citation comparison, Switzerland performs better in citations received than either Finland or Mexico. In both technologies examined, Finland and Mexico had similar attributes. The very similar attributes in Films, and the substantial difference in median and average citations received by the most highly cited Gene papers, show the necessity for comparing countries at the critical technology level rather than at the aggregate level (King, 2004).

APPENDIX 3 – TAXONOMY CLUSTERS – 200

1. Physical, Informational, Environmental and Economic Sciences

1.1. Physics, Information Science, and Applied Science

1.1.1. Economic Studies, Information Technology, Theoretical Mathematics

1.1.1.1. Economics / Sociology / Economic Models

- Cluster 24 (42) [firm 58.5%, wast 3.9%, market 2.2%, relationship 1.5%, knowledg 0.6%, compani 0.6%, environment 0.5%, oper 0.5%, mnc 0.5%, product 0.5%, multin 0.5%, paper 0.4%, finnish 0.4%, renew 0.4%, portfolio 0.4%] Focused on factors effecting the operation of Finnish businesses and companies, particularly new markets, establishment of multi-national corporations (MNCs), and growing environmental complexities.
- Cluster 134 (58) [market 27.4%, price 11.7%, countri 4.3%, economi 2.3%, seller 2.3%, buyer 1.5%, paper 1.1%, auction 1.1%, nordic 0.9%, run 0.6%, equilibrium 0.6%, labour 0.5%, nordic.countries 0.5%, effici 0.5%, product 0.5%] Focused on the state of the economy.
- Cluster 167 (65) [compani 11.7%, custom 8.1%, cost 5.6%, bank 4.2%, product 4.0%, capit 3.6%, road 2.7%, econom 1.6%, profit 1.6%, effici 1.4%, oper 1.2%, power 1.2%, market 1.1%, paper 0.9%, output 0.8%] Focused on economic trends and economic strategies.
- Cluster 193 (90) [articl 13.6%, polit 5.5%, parti 4.7%, vote 2.1%, countri 2.0%, social 1.4%, welfar 1.3%, cultur 1.3%, journal 1.2%, metaphor 1.1%, new 1.1%, finnish 1.0%, european 1.0%, stori 0.9%, state 0.9%] Focused on social, economic, and political aspects of Finland.
- Cluster 199 (108) [knowledg 4.2%, theori 3.5%, conceptu 3.4%, model 2.8%, social 2.0%, truth 1.7%, resourc 1.7%, practic 1.7%, concept 1.7%, framework 1.4%, theoret 1.3%, perspect 1.2%, natur 1.2%, rule 1.2%, search 1.0%] Focused on sociological concepts and theories.
- Cluster 38 (41) [learn 44.9%, instruct 2.2%, interact 1.7%, concept 1.7%, collabor 1.5%, learning.environments 1.2%, environ 1.1%, computer.supported 1.1%, inquiri 1.0%, design 1.0%, inform 0.9%, cours 0.8%, goal 0.8%, comput 0.6%, team 0.6%] Focused on the use of information technology in the educational process.
- Cluster 124 (45) [decis 22.8%, inform 14.3%, plan 3.5%, adopt 1.8%, uncertainti 1.5%, maker 1.3%, decision.maker 1.0%, smaa 1.0%, criteria 1.0%, innov 1.0%, swot 0.8%, data 0.8%, technolog 0.7%,

expert 0.7%, paper 0.7%] Focused on decision making processes (such as analytic hierarchy process (ahp)) and decision support systems.

- Cluster 139 (49) [project 27.2%, accid 3.7%, programm 2.8%, organis 2.2%, refer 1.1%, european 1.0%, panel 0.9%, design 0.9%, concept 0.8%, exposur 0.8%, construct 0.8%, indic 0.7%, fusion 0.7%, inform 0.7%, materi 0.6%] Focused on core recovery accident management research efforts / programs.

1.1.1.2. Information Technology / Communication Systems

- Cluster 56 (90) [softwar 51.6%, architectur 8.1%, tool 2.2%, product 2.0%, system 1.8%, paper 1.7%, model 1.0%, design 0.9%, compon 0.9%, software.product 0.7%, formal 0.7%, software.architecture 0.6%, applic 0.6%, product.line 0.6%, framework 0.5%] Focused on software engineering, and software development methods / processes.
- Cluster 185 (68) [system 15.2%, game 8.7%, paper 5.9%, machin 3.2%, agent 2.7%, applic 2.4%, real 1.3%, paper.machine 1.2%, player 1.2%, design 1.1%, real.time 1.1%, adapt 1.0%, control 1.0%, model 0.9%, oper 0.9%] Focused on (video) games designed for wireless networks and mobile communication systems.
- Cluster 115 (67) [mobil 26.6%, devic 5.0%, commun 4.9%, web 3.6%, messag 3.4%, servic 3.0%, inform 1.5%, applic 1.4%, environ 1.4%, memori 1.2%, agent 1.1%, person 1.0%, paper 0.8%, awar 0.8%, phone 0.8%] Focuses on mobile computing, communication system, and fusing sensory applications.
- Cluster 117 (79) [network 33.2%, wireless 8.8%, protocol 3.8%, secur 2.8%, internet 2.4%, mobil 1.7%, access 1.6%, system 1.5%, servic 1.3%, commun 1.2%, architectur 1.2%, traffic 1.1%, applic 0.9%, node 0.9%, paper 0.9%] Focuses on wireless networks and mobile communication devices, with emphasis on protocols and security.

1.1.2. Theoretical Mathematics and Applied Physics

1.1.2.1. Theoretical Mathematics / Signal Processing

- Cluster 27 (27) [code 57.4%, equal 3.7%, bound 1.1%, optim 1.0%, length 0.8%, equival 0.8%, codes.length 0.8%, binari 0.7%, alphabet 0.6%, equal.equal 0.6%, sequenc 0.5%, set 0.4%, pictur 0.4%, subset 0.4%] Focused on mathematical algorithms and linear codes.

- Cluster 84 (65) [channel 14.1%, receiv 4.3%, decod 3.1%, bit 2.9%, error 2.8%, fade 2.3%, cdma 2.1%, code 1.7%, estim 1.6%, interfer 1.6%, multipl 1.4%, algorithm 1.4%, ber 1.2%, multiple.access 1.2%, system 1.1%] Focused on communication technologies, particularly cellular communication network receivers and decoders, channel estimation processes, and channel interference problems.
- Cluster 122 (131) [algorithm 64.8%, comput 0.9%, learn 0.7%, optim 0.5%, paper 0.5%, linear 0.4%, network 0.4%, data 0.4%, effici 0.4%, approxim 0.4%, imag 0.4%, model 0.4%, iter 0.4%, time 0.4%, bit 0.3%] Focused on algorithms for various applications.
- Cluster 146 (58) [optim 12.0%, algorithm 8.0%, search 7.5%, text 3.0%, learn 2.4%, differential.evolution 2.3%, design 2.2%, topic 2.2%, approxim 1.3%, pattern 1.0%, paper 1.0%, effici 1.0%, solut 1.0%, represent 0.9%, set 0.9%] Focused on optimization algorithms, and signal processing methods such as finite impulse response (FIR).
- Cluster 161 (66) [finit 6.8%, automata 4.7%, languag 4.6%, algebra 4.3%, logic 3.9%, program 2.6%, theori 2.5%, class 2.4%, rule 1.8%, set 1.7%, function 1.7%, automaton 1.7%, finite.automata 1.6%, translat 1.6%, undecid 0.9%] Focused on the finite automata model of computation, and the theory of formal languages.
- Cluster 109 (43) [omega 14.3%, equal 7.5%, set 4.3%, equal.equal 2.8%, element 2.3%, omega.omega 2.3%, relat 2.1%, lambda 1.6%, cardin 1.6%, sum.rules 1.3%, isomorph 1.3%, sum 1.3%, subset 1.0%, kramers.kronig 1.0%, extens 0.9%] Focused on mathematical models.
- Cluster 150 (81) [space 13.3%, inequ 6.3%, theorem 4.8%, boundari 3.6%, map 3.1%, sobolev 3.0%, infin 2.0%, prove 1.5%, banach 1.5%, oper 1.3%, dirichlet 1.3%, integr 1.2%, function 1.2%, distort 1.2%, bound 1.2%] Focused on nonlinear mathematics and sobolev spaces.
- Cluster 149 (83) [equat 25.4%, solut 6.1%, boundari 3.2%, integr 2.2%, function 2.1%, oper 2.0%, element 2.0%, finit 1.6%, singular 1.4%, differenti 1.0%, condit 1.0%, numer 1.0%, gener 1.0%, ellipt 0.7%, linear 0.7%] Focused on solutions to differential equations and their use in practical applications.
- Cluster 164 (63) [perturb 5.2%, phase 3.9%, theori 3.7%, state 2.6%, vacuum 2.4%, observ 2.0%, cosmolog 1.9%, gaug 1.6%, equat 1.5%, baryon 1.2%, quark 1.1%, calcul 0.9%, flat 0.9%, qcd 0.9%, quantum 0.9%] Focused on the theory of phase transitions and critical phenomena.

- Cluster 170 (61) [nonlinear 11.3%, model 5.9%, nois 4.0%, speech 3.1%, dynam 2.8%, optim 2.4%, linear 2.3%, system 1.9%, time 1.6%, discret 1.4%, discrete.time 1.3%, continuous.time 1.2%, network 1.1%, stationari 1.0%, class 1.0%] Focused on nonlinear dynamic models, with emphasis on noise and speech (recognition).
- Cluster 189 (109) [model 45.0%, data 1.8%, uncertainti 1.6%, set 1.1%, simul 0.8%, moistur 0.6%, distribut 0.6%, paramet 0.6%, regress 0.6%, statist 0.5%, linear 0.5%, paper 0.4%, structur 0.4%, dynam 0.4%, nois 0.4%] Focused on the use of models to simulate biological and physical phenomena.
- Cluster 198 (102) [estim 6.5%, test 5.7%, data 3.0%, error 2.6%, statist 2.0%, simul 1.8%, distribut 1.7%, fuzzi 1.6%, covari 1.5%, model 1.5%, asymptot 1.3%, multivari 1.3%, covariance.matrix 0.9%, trait 0.9%, shift 0.9%] Focused on methods for analyzing data sets.
- Cluster 126 (56) [som 7.8%, color 5.7%, textur 4.6%, map 4.3%, self.organizing 2.8%, featur 2.4%, self.organizing.map 2.4%, organizing.map 2.4%, cluster 2.4%, visual 1.9%, classifi 1.9%, face 1.7%, histogram 1.6%, recognit 1.6%, self 1.4%] Focused on visualization algorithms such as self organizing map.
- Cluster 118 (72) [imag 52.9%, spot 1.8%, segment 1.0%, pixel 1.0%, compress 0.8%, digit 0.5%, visual 0.5%, inform 0.5%, microarrai 0.5%, paper 0.5%, model 0.5%, classif 0.4%, volum 0.3%, retriev 0.3%, new 0.3%] Focused on the physics of (electronic) imaging systems.
- Cluster 171 (81) [filter 25.8%, nois 8.1%, signal 5.0%, digit 2.7%, polynomi 1.6%, frequenc 1.2%, ey 0.9%, imag 0.9%, linear 0.9%, eeg 0.8%, kalman 0.7%, circuit 0.6%, time 0.6%, process 0.6%, cmo 0.5%] Focused on signal processing, signal filtering techniques and noise reduction.
- Cluster 183 (83) [antenna 8.5%, frequenc 5.6%, calibr 5.5%, ghz 4.5%, nois 2.4%, cabl 2.0%, reson 1.9%, voltag 1.7%, silicon 1.6%, arrai 1.5%, radiat 1.4%, measur 1.1%, devic 1.1%, radio 1.0%] Focused on communications equipment (particularly antennas), antenna calibration, and communication frequencies.
- Cluster 72 (63) [laser 30.6%, optic 3.6%, weld 3.3%, puls 2.6%, diod 2.2%, caviti 2.0%, semiconductor 1.8%, fiber 1.2%, mirror 1.2%, lock 1.1%, power 0.9%, gaa 0.9%, pump 0.7%, mode 0.7%, fiber.laser 0.7%] Focused on laser technologies, particularly laser welding, and vertical cavity surface emitting laser for applications such as high speed

optical interconnects, parallel data links, optical recording, 2-D scanning, and optical signal processing.

- Cluster 180 (79) [optic 17.8%, beam 4.9%, wavelength 4.3%, polar 3.3%, spectral 3.2%, sensor 2.8%, diffract 1.5%, america 1.5%, measur 1.1%, light 0.8%, lithographi 0.8%, filter 0.7%, fiber 0.7%, radio 0.7%, harmon 0.6%] Focused on use of laser beam devices such as diffractive optical elements.
- Cluster 101 (60) [galaxi 25.0%, bar 3.3%, luminos 2.3%, star 2.1%, sourc 1.6%, mass 1.5%, cloud 1.3%, galact 1.2%, redshift 1.2%, object 1.2%, mum 1.0%, emiss 1.0%, spiral 0.9%, optic 0.9%, line 0.9%] Focused on astronomy, particularly luminous stars and active galactic nuclei.
- Cluster 113 (46) [orbit 11.5%, rotat 5.5%, asteroid 5.2%, lunar 4.9%, rai 2.2%, solar 2.0%, period 2.0%, lightcurv 1.5%, earth 1.3%, mission 1.2%, instrument 1.0%, cix 1.0%, accret 0.9%, itokawa 0.8%, neptun 0.8%] Focused on asteroid orbit and topography, and lunar orbit and topograpy.
- Cluster 100 (46) [veloc 28.3%, radar 2.9%, crust 2.6%, profil 1.9%, stare 1.6%, seismic 1.5%, depth 1.3%, pulsat 1.2%, mantl 1.0%, flow 0.9%, telescop 0.9%, crustal 0.8%, radial 0.8%, turbul 0.7%, model 0.7%] Focused on geological instrumentation, and velocity measurement techniques using radar.
- Cluster 182 (59) [surfac 5.3%, stress 4.9%, contact 3.4%, disloc 2.4%, simul 2.0%, deform 1.8%, strain 1.5%, properti 1.3%, forc 1.2%, crack 1.1%, element 1.0%, finit 1.0%, etch 0.9%, elast 0.8%] Focused on the mechanics, manufacture, and use of structural and nonstructural materials in a variety of applications.
- Cluster 22 (32) [cartilag 17.7%, fdtd 6.8%, adi 3.7%, time.domain 3.2%, imped 2.5%, finite.difference.time 2.4%, difference.time.domain 2.4%, difference.time 2.4%, finite.difference 2.3%, domain 1.5%, modulu 1.5%, finit 1.3%, numer 1.3%, indent 1.2%, articular 1.2%] Focused on the use of finite difference time domain (fdtd) methods to model / analyze mechanical properties of materials and macromolecular structures and articular cartilage.
- Cluster 191 (88) [scatter 8.6%, model 7.9%, heat 2.0%, particl 2.0%, simul 2.0%, numer 1.8%, diffus 1.5%, ic 1.2%, approxim 0.9%, flow 0.9%, wind 0.9%, turbul 0.9%, inclus 0.8%, transfer 0.8%, surfac 0.8%, shape 0.7%] Focused on processes / reactions that create radiative

heat transfer as a byproduct, and single-scattering modeling for radiative flux and radiance calculations.

- Cluster 81 (59) [wave 46.0%, plane 1.7%, slab 1.6%, propag 1.6%, electromagnet 0.9%, surfac 0.8%, arrai 0.8%, numer 0.7%, plane.wave 0.7%, field 0.7%, particl 0.6%, shock 0.6%, backward 0.6%, direct 0.5%, frequenc 0.5%] Focuses on waves and spatial dispersion.
- Cluster 51 (48) [solar 24.3%, sunspot 4.6%, magnet 3.3%, solar.wind 2.5%, cycl 2.5%, ena 2.1%, wind 2.0%, solar.activity 1.6%, period 1.1%, barkhausen.noise 1.1%, barkhausen 1.1%, activ 1.1%, cosmic.ray 1.1%, flare 1.1%, heliospher 1.0%] Focused on solar activity and interplanetary magnetic fields.
- Cluster 35 (53) [auror 9.6%, ionospher 6.3%, arc 5.7%, current 4.6%, field 3.4%, altitud 2.7%, magnetospher 2.6%, substorm 2.2%, imf 2.1%, magnet 1.7%, region 1.2%, satellit 1.2%, fac 1.1%, geomagnet 1.1%, electrojet 1.1%] Focused on the ionosphere, interplanetary magnetic fields and solar winds, particularly aurora borealis phenomena.
- Cluster 114 (111) [magnet 38.3%, field 12.6%, magnetic.field 9.8%, temperatur 1.4%, martensit 0.8%, msm 0.6%, alloi 0.5%, current 0.5%, strain 0.4%, magnetic.fields 0.4%, rotor 0.4%, field.induced 0.4%, polar 0.4%, spin 0.4%, measur 0.3%] Focused on magnetic fields and materials, particularly, magnetic shape memory alloys.
- Cluster 130 (44) [field 26.6%, coher 6.0%, polar 3.0%, scalar 1.7%, electr 1.5%, electromagnet 1.4%, degree.coherence 1.3%, electric.field 1.0%, degre 0.7%, dep 0.7%, partially.coherent 0.6%, azimuth 0.5%, partial 0.5%, model 0.5%, dipol 0.5%,] Focused on the propagation and retrieval of cross-spectral density.
- Cluster 59 (43) [current 33.1%, tape 5.0%, voltag 4.7%, critical.current 1.8%, field 1.8%, magnet 1.7%, measur 1.4%, self.field 1.0%, voltage.current 0.9%, critic 0.9%, transistor 0.8%, densiti 0.8%, current.density 0.8%, sheath 0.6%, dissip 0.5%] Focused on methods and processes for measuring magnetic and electric fields.

1.1.2.2. Applied Physics

- Cluster 61 (55) [mode 10.4%, elm 5.4%, plasma 5.0%, jet 3.5%, gyrotron 3.3%, heat 2.9%, edg 2.8%, ion 2.6%, icrf 1.7%, elmi 1.5%, oscil 1.4%, mhd 1.1%, power 1.0%, profil 1.0%, stabil 1.0%] Focused on plasma physics and gyrotron systems.
- Cluster 37 (51) [vortex 26.0%, superfluid 7.2%, bose 6.2%, vortic 5.5%, condens 4.3%, bose.einstein 2.2%, einstein 2.0%, fermi 1.2%, turbul

1.1%, atom 0.9%, quantiz 0.9%, rotat 0.9%, temperatur 0.7%, trap 0.7%, bose.einstein.condensate 0.7%] Focused on gaseous superfluid phases, vortices, and bose einstein condensates.

- Cluster 129 (54) [tunnel 7.7%, electron 6.3%, coupl 5.1%, junction 3.9%, superconductor 3.5%, josephson 2.8%, temperatur 2.3%, transistor 1.8%, superconduct 1.7%, normal.metal 1.3%, charg 1.2%, energi 1.2%, cooper.pair 1.1%, structur 0.9%, supercurr 0.9%] Focused on josephson junctions, joshephson coupling, tunneling and superconducting metals.
- Cluster 120 (79) [quantum 12.1%, spin 10.6%, electron 7.1%, dot 5.1%, densiti 2.6%, state 2.2%, quantum.dot 1.6%, theori 1.4%, hole 1.4%, density.functional 1.4%, noncommut 1.3%, functional.theory 1.1%, density.functional.theory 1.1%, magnet 0.9%, calcul 0.9%] Focused on physics used for calculating the electronic structure, properties of matter with methods such as spin deinsity functional theory, and analyzing the spectra of quantum dots.
- Cluster 177 (77) [simul 7.0%, cluster 6.2%, ion 5.4%, energi 3.4%, dynam 1.9%, molecular 1.7%, monte.carlo 1.5%, model 1.5%, carlo 1.5%, mont 1.5%, molecular.dynamics 1.5%, interact 1.1%, potenti 1.1%, calcul 1.1%, atom 0.9%] Focused on molecular dynamic simulations, and use of monte carlo methods for evaluating nanocluster.
- Cluster 168 (98) [energi 48.6%, calcul 1.2%, surfac 1.2%, atom 0.8%, interact 0.7%, potential.energy 0.7%, system 0.7%, structur 0.6%, state 0.6%, electron 0.5%, potenti 0.5%, model 0.5%, adsorpt 0.5%, conform 0.5%, densiti 0.5%] Focused molecular quantum mecanics, particularly on potential energy surface (PES).
- Cluster 39 (41) [auger 9.2%, spectra 5.3%, energi 4.7%, electron 4.4%, excit 4.2%, arp 3.8%, photon 2.1%, state 1.8%, photoelectron 1.7%, reson 1.1%, spectrum 1.1%, angular 1.0%, photon.energy 0.8%, core 0.8%, resolut 0.8%] Focused on experimental tools / methods for spectroscopy, particularly Auger electron spectroscopy.
- Cluster 128 (66) [state 10.7%, excit 9.4%, band 8.3%, vibrat 8.3%, absorpt 2.2%, spectra 1.5%, spectrum 1.2%, stretch 1.2%, excited.state 1.2%, shift 1.0%, rotat 0.9%, reson 0.9%, energi 0.8%, dft 0.8%, molecul 0.7%] Focused on the approach for describing the ground state properties of metals, semiconductors, and insulators.
- Cluster 70 (69) [decai 27.0%, gamma 4.4%, state 4.1%, beta 2.1%, recoil 2.0%, alpha 1.5%, nuclei 1.5%, rai 1.1%, gamma.ray 1.1%, beta.decay 1.1%, detector 1.0%, proton 1.0%, separ 1.0%, gas.filled

0.9%, alpha.decay 0.9%] Focused on experimental and theoretical physics, particularly gamma-ray decay.

- Cluster 92 (76) [bar 8.0%, gev 5.4%, hadron 4.7%, mass 4.0%, decai 3.6%, quark 2.9%, boson 2.4%, higg 2.3%, delphi 2.2%, standard.model 1.7%, collis 1.6%, meson 1.5%, lepton 1.3%, supersymmetr 1.0%, collid 1.0%, energi 1.0%, syst 1.0%, search 1.0%, model 0.9%, stat 0.9%, mass.energies 0.9%] Focused on theoretical models used in particle physics, with emphasis on hadrons, quark masses, and bosons.

1.2. Chemistry (Organic / Inorganic), Environment, and Ecology

1.2.1. Material Science, Organic Chemistry, Environmental and Ecological Studies

1.2.1.1. Physical Chemistry / Metallurgy / Material Science

- Cluster 6 (30) [vacanc 25.0%, defect 9.6%, positron 7.6%, annihil 4.1%, positron.annihilation 3.3%, gan 3.3%, anneal 2.4%, dope 1.7%, annihilation.spectroscopy 1.5%, positron.annihilation.spectroscopy 1.5%, grown 1.0%, complex 0.9%, irradi 0.9%, epitaxi 0.8%, spectroscopi 0.7%] Focuses on experimental physics in material research, and analysis of vacancies / defects in material crystal lattices, with emphasis on positron annihilation.
- Cluster 103 (54) [anneal 10.7%, dope 6.3%, ion 4.1%, luminesc 2.8%, gaa 2.3%, interdiffus 2.3%, eu2 2.0%, layer 1.8%, implant 1.7%, temperatur 1.6%, epitaxi 1.6%, getter 1.2%, boron 1.2%, diffus 1.1%, quantum 1.1%] Focused on ion beam analysis techniques for quantitative analysis of light elements in solids, with emphasis on the luminescent properties of annealed or doped materials.
- Cluster 159 (81) [surfac 18.6%, atom 9.6%, layer 6.0%, silicon 2.5%, sputter 1.7%, structur 1.3%, microscopi 1.2%, tip 0.9%, irradi 0.7%, adsorpt 0.7%, cluster 0.7%, rough 0.7%, atomic.force 0.6%, electron 0.6%, forc 0.6%] Focused on methods, such as atomic force microscopy, for determining material surface structures at the atomic level.
- Cluster 99 (40) [alloi 16.0%, nanotub 13.0%, martensit 5.2%, copper 3.4%, solder 2.0%, interfac 1.6%, crystal 1.6%, structur 1.5%, temperatur 1.4%, layer 1.3%, carbon.nanotubes 1.3%, carbon 0.9%, electron 0.9%, monolay 0.9%, ion 0.8%] Focused on metallurgy, memory shaped alloys, and nanomaterials.

- Cluster 75 (44) [steel 25.1%, corros 6.1%, slag 5.2%, bubbl 1.9%, stainless 1.7%, entrain 1.4%, iron 1.4%, alloi 1.3%, ga 1.3%, temperatur 1.2%, heat 1.2%, sprayform 1.0%, surfac 0.9%, materi 0.8%, liquid 0.8%] Focused on metallurgy, with emphasis on gas bubbles resulting during the metal separation / slag removal process, and effects of alloying elements in the corrosion resistance of steels.
- Cluster 23 (34) [fiber 13.8%, frc 9.0%, strength 4.1%, composit 3.9%, specimen 3.2%, reinforc 2.7%, bond 2.5%, resin 2.3%, bond.strength 2.1%, flexur 1.9%, mpa 1.8%, materi 1.8%, fractur 1.7%, glass 1.3%, fibr 1.1%,] Focused on (fiber reinforced) composite materials.
- Cluster 93 (37) [bind 6.3%, structur 5.4%, residu 2.8%, conform 2.7%, active.site 2.7%, site 2.2%, enzym 1.9%, loop 1.8%, crystal 1.8%, ppase 1.5%, avidin 1.5%, bond 1.1%, glu 1.0%, beta 1.0%, cellulos 1.0%] Focused on biochemical reactions, particularly of metal binding sites.
- Cluster 54 (72) [coat 57.8%, surfac 2.3%, silica 1.5%, wear 1.0%, laser 0.6%, cure 0.6%, starch 0.6%, properti 0.5%, base.paper 0.5%, glass 0.5%, poros 0.5%, colour 0.4%, sprai 0.4%, paper 0.4%, print 0.4%] Focused on substrates and coatings, particularly silica, and their use in various applications.
- Cluster 31 (104) [film 62.2%, deposit 3.0%, thin 2.3%, thin.films 1.3%, layer 0.7%, temperatur 0.5%, substrat 0.5%, thin.film 0.5%, thick 0.4%, films.grown 0.4%, surfac 0.4%, dielectr 0.4%, ald 0.3%, atom 0.3%, properti 0.3%] Focused on deposition and growth of thin film materials for microelectronics applications.
- Cluster 123 (48) [deposit 10.8%, ink 7.1%, print 4.4%, layer 4.1%, film 3.1%, ald 2.4%, oxid 2.3%, electrod 1.7%, multilay 1.6%, atomic.layer 1.5%, layer.deposition 1.3%, thick 1.2%, atomic.layer.deposition 1.1%, mum 1.1%, dielectr 1.1%] Focused on material science of thin films, microelectronics, processes such as atomic layer deposition, and ink resins for printing.

1.2.1.2. Organic Chemistry

- Cluster 60 (92) [catalyst 47.0%, hydrogen 5.5%, reaction 2.0%, activ 1.6%, mao 1.6%, alumina 1.3%, catalyt 1.2%, zeolit 1.0%, support 0.9%, adsorpt 0.8%, isomer 0.7%, acid 0.6%, xylen 0.6%, silica 0.5%, deactiv 0.5%] Focused on the use of selective catalytic hydrogen in chemical reactions.

- Cluster 178 (79) [reaction 23.1%, enantioselect 3.6%, kinet 2.0%, hydrogen 1.9%, adduct 1.3%, butano 1.3%, acid 1.1%, acyl 1.0%, lipas 1.0%, model 0.9%, product 0.9%, reactor 0.8%, oxid 0.7%, solvent 0.7%, catalyst 0.6%] Focused on organic reactions that allow chemical transformations.
- Cluster 166 (54) [compound 4.9%, synthesi 4.8%, deriv 3.2%, acid 2.4%, bond 2.4%, bi 2.2%, protect 2.1%, molecul 2.0%, cleavag 1.8%, analogu 1.7%, yield 1.7%, synthes 1.2%, substitut 1.1%, reaction 1.0%, linker 0.9%] Focused on chemical reactions, synthesis of compounds, and analyzing the molecular mechanics and dynamics of various substances.
- Cluster 138 (61) [nmr 10.5%, methyl 5.4%, substitu 4.2%, ring 3.8%, compound 3.3%, structur 2.2%, aryl 2.0%, phenyl 1.7%, reaction 1.5%, isom 1.0%, electron 1.0%, bond 1.0%, tautomer 0.9%, chloroform 0.8%, chemic 0.8%] Focused on nuclear magnetic resonance (NMR) spectroscopy, .
- Cluster 160 (107) [complex 13.9%, ligand 8.4%, bond 2.7%, ion 2.4%, coordin 2.3%, reaction 2.1%, cation 2.0%, atom 1.7%, structur 1.7%, anion 1.7%, bridg 1.6%, metal 1.4%, tran 1.2%, octahedr 1.1%, compound 1.0%] Focused on chemical reactions of various compounds, particularly ligand exchange (substitution) reactions involving complex metal ions.
- Cluster 14 (35) [angstrom 35.7%, crystal 12.0%, angstrom.beta 2.0%, space.group 2.0%, beta 1.6%, monoclin 1.5%, degre 1.0%, unit.cell.parameters 0.9%, cell.parameters 0.9%, unit.cell 0.9%, diffract 0.8%, ref 0.8%, space 0.7%, cage 0.7%, structur 0.7%] Focused on various material microstructures and crystal lattices, with an emphasis unit cell parameters.
- Cluster 65 (42) [poli 8.7%, polym 8.3%, chain 4.9%, methacryl 3.6%, copolym 3.4%, block 2.4%, polymer 1.9%, supramolecul 1.7%, p4vp 1.4%, pvp 1.4%, scatter 1.2%, complex 0.8%, oligom 0.8%, structur 0.8%, solut 0.8%] Focused on synthesis of polymer / copolymer materials.
- Cluster 88 (42) [membran 28.7%, bilay 3.1%, lipid 3.0%, partit 1.2%, dhsm 1.1%, water 1.0%, liquid 0.9%, foul 0.9%, drug 0.8%, acid 0.8%, monolay 0.8%, c12e8 0.8%, pani 0.7%, interfac 0.7%, sphingomyelin 0.7%] Focused on chemical reaction of various compounds, especially reactions involving membranes and lipid bilayers.

- Cluster 156 (54) [solut 12.0%, ion 7.8%, exchang 4.2%, sodium 3.3%, surfact 2.5%, acid 1.7%, calcium 1.5%, aqueou 1.4%, water 1.4%, carbon 1.3%, ion.exchange 1.2%, phase 1.1%, electrod 0.9%, concentr 0.7%, select 0.7%] Focused on chemical reactions involving ionic solutions / sodium chloride solutions.
- Cluster 186 (64) [acid 18.1%, water 7.8%, sulfur 2.1%, solubl 1.5%, sulfuric.acid 1.5%, lactos 1.3%, juic 1.2%, xylanas 1.1%, deproton 0.9%, chelat 0.9%, separ 0.8%, solut 0.7%, medium 0.7%, bromid 0.7%, capsul 0.6%] Focused on water purification and water treatment, with emphasis on the effects of sulfuric acid on water.
- Cluster 192 (98) [compound 8.7%, sampl 3.4%, mass 3.0%, spectrometri 2.2%, chromatographi 1.9%, liquid 1.7%, esi 1.7%, detect 1.7%, mass.spectrometry 1.4%, chromatograph 1.3%, lignan 1.3%, extract 1.2%, ioniz 1.2%, ga 1.0%, phase 0.9%] Focused on analytical methods of chemistry such as liquid chromatography and mass spectrometry.
- Cluster 71 (44) [extract 28.3%, pah 5.2%, compound 2.6%, liquid 1.6%, water 1.4%, pyren 1.1%, anthracen 0.9%, phwe 0.9%, sfe 0.8%, bap 0.7%, polycyclic.aromatic 0.7%, polycycl 0.7%, hydrocarbon 0.7%, polycyclic.aromatic.hydrocarbons 0.6%, sediment 0.6%] Focused on biochemistry, with emphasis on the extraction of polycyclic aromatic hydrocarbon (PAH) compounds.
- Cluster 76 (39) [oxid 13.9%, iron 7.6%, antioxid 6.2%, phenol 4.1%, extract 3.1%, scaveng 3.0%, radic 2.7%, radical.scavenging 2.4%, acid 1.9%, compound 1.7%, oil 1.6%, fbr 1.5%, iron.oxidation 1.2%, rapese 1.1%, tlc 0.8%] Focused on oxidation reactions and iron-oxidation.
- Cluster 140 (48) [remov 7.0%, sludg 6.7%, chemic 2.2%, wastewat 2.2%, digest 1.7%, wast 1.7%, dehp 1.6%, water 1.6%, treatment 1.6%, compost 1.5%, filtrat 1.5%, biodegrad 1.4%, cod 1.3%, sulphat 1.3%, organ 1.2%] Focused on wastewater treatment, especially sludge removal.
- Cluster 107 (58) [lignin 17.9%, wood 5.7%, fiber 4.2%, cellulos 3.4%, extract 2.6%, laccas 2.3%, resin 2.0%, liquor 1.9%, treatment 1.2%, pulp 1.1%, chip 1.1%, water 1.0%, cellulas 1.0%, degrad 0.9%, heat 0.9%] Focused on the conversion of wood into pulp, chips, and other products, as well as the extraction of lignin from wood.
- Cluster 28 (79) [pulp 49.8%, bleach 4.6%, kraft 3.6%, fibr 2.7%, lignin 2.2%, delignif 1.7%, kraft.pulp 1.4%, chemic 1.3%, mill 1.2%, cook

1.1%, softwood 0.6%, wash 0.6%, paper 0.5%, oxygen 0.5%, extract 0.5%] Focused on pulp and paper production technologies.

1.2.2. Environmental and Ecological Studies

1.2.2.1. Environmental Studies

- Cluster 152 (78) [particl 37.3%, size 4.7%, nanoparticl 2.7%, particle.size 2.5%, floccul 0.8%, surfac 0.7%, carbon 0.7%, zeta 0.6%, distribut 0.6%, diamet 0.6%, pore 0.5%, suspens 0.5%, nickel 0.5%, spheric 0.5%, deposit 0.5%] Focused on issues related to nanoparticles (size, stability, etc.) in various applications, such as drugs and structural materials.
- Cluster 29 (70) [particl 24.2%, nucleat 14.8%, aerosol 8.8%, condens 2.3%, particle.formation 1.9%, concentr 1.7%, size 1.6%, format 1.4%, growth 1.2%, mode 1.2%, new.particle 0.8%, coastal 0.8%, new.particle.formation 0.7%, droplet 0.7%, vapor 0.6%, acid 0.6%] Focused on particle nucleation and aerosols.
- Cluster 91 (39) [dust 7.0%, fuel 6.5%, boiler 6.1%, bed 4.3%, combust 4.2%, indoor 3.7%, particl 1.9%, coal 1.8%, emiss 1.6%, fluidiz 1.6%, flue 1.5%, build 1.1%, concentr 1.1%, ash 1.0%, air 1.0%] Focused on air quality and how it is effected by construction, power generation, particularly fluidized bed combustion, coal ash / dust and fuel emissions.
- Cluster 155 (44) [dry 8.1%, heat 8.0%, water 4.9%, temperatur 3.1%, diffus 2.3%, air 1.7%, tube 1.1%, flake 1.1%, transfer 1.1%, heat.transfer 1.0%, measur 1.0%, exchang 0.9%, theophyllin 0.8%, fluid 0.7%, pigeon 0.7%] Focused on heat (mass) transfer.
- Cluster 175 (62) [temperatur 33.0%, degreesc 4.7%, facet 1.1%, strain 0.9%, melt 0.8%, gel 0.8%, starch 0.8%, pvcl 0.8%, ca2 0.7%, oxid 0.7%, isotact 0.6%, crystal 0.5%, fraction 0.5%, growth 0.4%, 7degreesc 0.4%] Focused on the thermodynamics of phase transitions in various chemical reactions.
- Cluster 47 (36) [rock 8.2%, granit 6.3%, or 4.6%, miner 3.8%, pge 3.6%, grain 2.4%, magma 1.9%, mafic 1.8%, intrus 1.7%, zone 1.6%, quartz 1.5%, sulfid 1.4%, magmat 1.2%, melt 1.2%, dike 1.2%] Focused on mining, granite rocks, and platinum group elements (PGE).
- Cluster 188 (63) [climat 3.3%, deposit 3.0%, precipit 2.0%, finland 1.8%, north 1.7%, area 1.7%, palsa 1.4%, scenario 1.3%, northern 1.1%, zone 1.1%, temperatur 1.1%, region 1.1%, sector 0.9%, concentr 0.9%, winter 0.8%] Focused on the effects of regional and global climate changes on Finland.

- Cluster 9 (40) [ic 59.8%, sea 2.7%, snow 2.2%, sea.ice 2.0%, glacier 1.1%, svalbard 1.1%, ice.core 0.9%, core 0.7%, surfac 0.6%, water 0.5%, accumul 0.4%, thick 0.4%, centuri 0.4%, cpp 0.4%, slip 0.3%, ice.thickness 0.3%, blue.ice 0.3%, sheet 0.3%, baltic 0.3%] Focuses on ice core research for identifying history and future climatic and environmental patterns.
- Cluster 127 (88) [sea 23.3%, baltic 11.4%, baltic.sea 7.3%, bloom 2.6%, gulf 2.0%, area 1.4%, her 1.2%, water 1.1%, gulf.finland 0.9%, phytoplankton 0.9%, temperatur 0.8%, speci 0.7%, biomass 0.6%, atlant 0.6%, salin 0.6%] Focused on the nutrient input, salinity, and other biological properties of the Baltic sea.
- Cluster 169 (62) [water 30.3%, contamin 1.8%, sampl 1.8%, diatom 1.8%, drinking.water 1.6%, stream 1.6%, drink 1.3%, concentr 1.1%, moss 0.9%, commun 0.9%, exposur 0.8%, environment 0.7%, emiss 0.6%, fen 0.5%, microbi 0.5%, mycobacteria 0.5%] Focused on water contamination, water treatment.
- Cluster 34 (37) [sediment 40.6%, nutrient 4.1%, water 2.8%, alga 2.4%, concentr 1.8%, basin 1.2%, groundwat 1.0%, oligochaet 0.8%, contamin 0.8%, variegatu 0.7%, deposit 0.6%, tcbp 0.6%, pore.water 0.6%, denitrif 0.6%, bioaccumul 0.5%] Focused on agriculture, and the analysis of sediments and nutrients in water..
- Cluster 33 (63) [lake 56.5%, water 3.8%, ch4 1.7%, sediment 1.6%, eutroph 1.3%, deposit 0.7%, littor 0.7%, diatom 0.5%, aquat 0.5%, catchment 0.5%, phosphoru 0.4%, delta 0.4%, nutrient 0.4%, basin 0.4%, emiss 0.4%] Focused on the effects of emissions on the Finnish ecosystem, especially lakes and other bodies of water.
- Cluster 96 (53) [ch4 28.3%, co2 5.4%, emiss 4.8%, concentr 2.6%, carbon 2.1%, flux 1.5%, season 1.4%, soil 1.3%, peatland 1.0%, ch4.emissions 0.9%, plasterboard 0.9%, methan 0.8%, wetland 0.7%, ch4.oxidation 0.7%, atmospher 0.7%] Focused on the effect of CH4 and CO2 emissions on the environment.
- Cluster 18 (32) [milk 29.8%, cow 16.6%, calv 4.0%, feed 3.0%, lactat 1.8%, lamb 1.7%, concentr 1.6%, breast 1.1%, silag 1.0%, ferment 0.7%, dairi 0.7%, breast.feeding 0.6%, yield 0.6%, milk.yield 0.5%, loose.housing 0.5%] Focused on the production of milk / dairy products.
- Cluster 87 (145) [soil 60.9%, miner 1.3%, humu 1.3%, concentr 0.9%, forest 0.9%, site 0.7%, horizon 0.5%, ash 0.5%, pine 0.5%, seedl 0.5%, organ 0.5%, microbi 0.4%, mineral.soil 0.4%, water 0.4%, fertil

0.4%] Focused on forestry and the effects of soil properties in the growth of particular tree species.

- Cluster 62 (43) [forest 19.2%, carbon 6.7%, climat 4.5%, flux 4.1%, tax 4.0%, sink 3.1%, ecosystem 1.9%, trade 1.8%, net 1.3%, litterfal 1.2%, aerosol 1.0%, soil 0.9%, eddi 0.8%, site 0.8%, model 0.8%] Focused on forestry, wood, and wood-based products, with an emphasis on carbon sinks and their effects on climate change.
- Cluster 40 (87) [forest 53.6%, stand 4.9%, speci 2.9%, tree 2.2%, landscap 1.2%, area 1.0%, site 1.0%, natur 0.9%, habitat 0.8%, plot 0.7%, inventori 0.6%, patch 0.4%, manag 0.4%, divers 0.4%, wood 0.4%] Focused on forestry, particularly forest / ecological planning, species diversification, and landscape management.
- Cluster 77 (76) [tree 26.3%, stand 5.7%, stem 4.4%, heartwood 3.8%, height 2.9%, growth 2.3%, diamet 2.2%, branch 2.2%, pine 2.1%, model 1.9%, knot 1.7%, crown 1.3%, spruce 1.2%, scot 1.0%, xylem 0.8%] Focused on forestry with particular emphasis on diameter distribution of stock, and models for (tree / forest) growth simulation.
- Cluster 48 (50) [needl 11.3%, pine 7.6%, scot 4.8%, scots.pine 4.8%, wood 4.3%, seed 3.7%, co2 2.9%, sylvestri 1.8%, pinu 1.6%, pinus.sylvestris 1.6%, concentr 1.5%, temperatur 1.5%, tree 1.4%, scots.pine.pinus 1.4%, pine.pinus 1.4%, pollen 1.4%, pine.pinus.sylvestris 1.3%, elev 1.3%, orchard 1.1%, climat 0.9%, fertil 0.9] Focused on forestry, particularly the scots pine.
- Cluster 15 (34) [ozon 39.9%, ecotyp 5.7%, dormanc 2.8%, stomat 2.6%, clone 1.2%, photoperiod 1.0%, bud 1.0%, stratospher 1.0%, stomatal.conductance 0.9%, chamber 0.9%, leaf 0.8%, photosynthesi 0.8%, tree 0.7%, leav 0.6%, net 0.6%] Focused on the effect of ozone exposure on the ecosystem.
- Cluster 147 (89) [plant 20.2%, seedl 11.2%, herbivor 2.4%, growth 2.2%, shoot 2.2%, root 2.0%, leav 1.6%, leaf 1.4%, shade 1.3%, birch 1.1%, defoli 1.0%, resist 1.0%, nutrient 1.0%, phenol 1.0%, cultivar 0.9%] Focused on particular species of trees, such as the birch betula pendula, and leaf damage due to (insect) herbivores.

1.2.2.2. Ecological Studies

- Cluster 165 (70) [popul 18.0%, breed 7.6%, size 3.2%, food 1.9%, seed 1.7%, clutch 1.6%, reproduct 1.3%, bird 0.9%, pollin 0.9%, number 0.9%, clutch.size 0.9%, trait 0.9%, individu 0.9%, genet 0.8%, brood

0.8%] Focused on various species (particularly birds), their reproductive habits and population trends.

- Cluster 136 (52) [genet 7.1%, femal 6.4%, twin 6.3%, reproduct 5.4%, offspr 5.0%, popul 4.6%, food 2.3%, sex 1.8%, brood 1.5%, breed 1.3%, fit 1.3%, trait 1.2%, parent 1.1%, surviv 0.9%, male 0.8%] Focused on genetic variability of species and breeding patterns.
- Cluster 78 (63) [male 39.7%, femal 11.4%, sexual 4.2%, mate 3.3%, sex 1.6%, reproduct 1.2%, nest 0.8%, males.females 0.7%, queen 0.7%, ag 0.6%, encapsul 0.6%, immun 0.6%, popul 0.6%, immunocompet 0.5%, egg 0.5%] Focused on reproductive issues for human, animals, and insects.
- Cluster 3 (17) [egg 50.0%, nest 10.3%, coregoni 1.9%, layout 1.8%, femal 1.7%, bug 1.3%, hatch 1.3%, lai 1.2%, nesting.layouts 1.0%, predat 1.0%, egg.size 0.9%, egg.laying 0.8%, chick 0.7%, bifilosa 0.6%] Focuses on the effect of microorganisms (e.g., argulus coregoni, Acartia bifilosa) on the reproductive behavior of fish.
- Cluster 19 (27) [river 38.9%, salmon 8.6%, fish 2.8%, grayl 1.9%, smolt 1.8%, pcb 1.5%, stock 1.2%, seneg 0.9%, migrat 0.9%, pcp 0.8%, water 0.8%, pbde 0.8%, atlantic.salmon 0.8%, atlant 0.6%, reservoir 0.5%, vegetation.cover 0.5%] Focused on the effect of water chemistries on fish populations, especially salmon.
- Cluster 85 (53) [fish 14.6%, stock 7.6%, whitefish 6.6%, lake 5.6%, vendac 3.4%, catch 2.5%, perch 2.0%, year.class 1.7%, fisheri 1.3%, coregonu 1.2%, trout 1.1%, gillnet 1.0%, pele 0.9%, class 0.8%, growth 0.8%] Focused on fishing industry and fish farms, with emphasis on stock, whitefish, and perch..
- Cluster 12 (46) [predat 32.3%, vole 18.9%, prei 9.1%, weasel 2.1%, popul 1.4%, speci 1.3%, nest 1.3%, habitat 1.0%, breed 1.0%, lem 0.9%, mink 0.9%, owl 0.7%, densiti 0.6%, bird 0.6%, vole.populations 0.5%] Focused on various predatory species.
- Cluster 68 (43) [habitat 29.1%, patch 9.9%, bird 4.8%, dispers 4.4%, landscap 1.7%, trap 1.2%, densiti 1.2%, stream 1.0%] Focused on wildlife habitats and migration patterns, particularly of birds.
- Cluster 163 (144) [speci 59.3%, habitat 1.3%, genu 0.7%, distribut 0.6%, taxa 0.5%, new 0.5%, graze 0.4%, morpholog 0.4%, plant 0.4%, lichen 0.4%, genera 0.4%, grassland 0.4%, abund 0.3%, species.new 0.3%, isol 0.3%] Focused on forest ecology research, particularly quantity and richness of various plant species.

- Cluster 20 (40) [host 36.7%, parasit 18.7%, speci 6.0%, larva 1.3%, fish 1.2%, grb 0.7%, plant 0.7%, parasitoid 0.7%, host.plant 0.6%, paranoplocephala 0.6%, butterfli 0.6%, beetl 0.5%, genu 0.5%, parasite.species 0.5%, popul 0.5%] Focused on ecological species, particularly host plants, parasites, larvae, fish..
- Cluster 131 (45) [pcr 17.3%, dna 4.4%, genotyp 2.5%, detect 2.0%, infect 2.0%, carotovora 1.7%, antigen 1.6%, cultur 1.4%, plant 1.3%, sampl 1.1%, virul 1.1%, assai 1.1%, viru 0.8%, b19 0.8%, gene 0.8%] Focused on the polymerase chain reaction for generating copies of fragmented DNA.
- Cluster 83 (42) [isol 36.0%, resist 4.4%, ampicillin 2.1%, outbreak 1.8%, coli 1.4%, pilosicoli 1.3%, ribotyp 1.3%, serotyp 1.0%, pfge 1.0%, lactamas 0.9%, beta.lactamase 0.9%, tuberculosi 0.9%, mgit 0.8%, antimicrobi 0.7%, ciprofloxacin 0.6%] Focused on treatment for various viral and bacterial diseases.
- Cluster 102 (91) [strain 54.3%, isol 3.0%, sequenc 1.7%, gene 1.5%, probiot 1.4%, serotyp 0.7%, adhes 0.6%, lactobacillu 0.6%, phylogenet 0.6%, genom 0.5%, viru 0.5%, 16 0.4%, bacteri 0.4%, xylos 0.4%, bacteria 0.4%] Focused on various bacteria strains, especially isolated gene sequences..

2. Biochemistry and Biology

2.1. Biochemistry, Biology, and Medicine

2.1.1. Biochemistry

2.1.1.1. Laboratory Biological Trials

- Cluster 55 (47) [viru 13.7%, rna 10.4%, protein 6.8%, viral 3.5%, virus 3.0%, infect 2.9%, genom 2.1%, strand 1.5%, sequenc 1.5%, pva 1.4%, replic 1.3%, capsid 1.1%, bacteriophag 0.9%, vp2 0.8%] Focused on the genetic makeup of various viruses.
- Cluster 158 (72) [protein 27.9%, gene 4.3%, amino 1.8%, acid 1.2%, laccas 1.1%, cerevisia 1.0%, sequenc 1.0%, proteinas 1.0%, mutant 1.0%, reesei 0.9%, encod 0.9%, secret 0.8%, amino.acid 0.7%, kda 0.7%, mitochondri 0.6%] Focused on genetic structure of proteins and amino acids.
- Cluster 10 (23) [nephrin 20.4%, pltp 8.5%, heat.shock 7.1%, shock 5.9%, heat 3.9%, protein 3.0%, hsf1 1.9%, hsp70 1.6%, express 1.6%, podocyt 1.1%, slit 0.9%, hsp 0.9%, stress 0.9%, cell 0.7%, nphs1 0.7%, dep 0.6%, diaphragm 0.6%, phospholipid.transfer 0.5%,

stress.granules 0.5%] Focuses on heat shock proteins and their role in cancer, especially nephrin and phosphor-lipid transfer proteins (PLTP).

- Cluster 95 (54) [domain 19.6%, bind 7.8%, protein 5.9%, pdi 2.5%, residu 2.5%, collagen 2.0%, site 1.6%, disulfid 1.5%, termin 1.4%, fold 1.3%, sequenc 0.9%, enzym 0.9%, erp57 0.9%, region 0.7%, famili 0.7%] Focused on binding proteins, biding domains, and their biochemical makeup.
- Cluster 93 (37) [bind 6.3%, structur 5.4%, residu 2.8%, conform 2.7%, active.site 2.7%, site 2.2%, enzym 1.9%, loop 1.8%, crystal 1.8%, ppase 1.5%, avidin 1.5%, bond 1.1%, glu 1.0%, beta 1.0%, cellulos 1.0%] Focused on biochemical reactions, particularly of metal binding sites.
- Cluster 116 (43) [peptid 8.7%, antibodi 5.0%, avidin 3.4%, amino 2.7%, mutant 2.6%, bind 2.5%, protein 2.5%, sulfonamid 2.2%, acid 2.1%, biotin 2.0%, mab 1.8%, mutat 1.3%, assai 1.3%, affin 1.3%, amino.acid 1.2%] Focuses biological synthesis and evaluation of peptide proteins for applications such as hormonal regulation and antibiotic activities.
- Cluster 172 (66) [bind 12.6%, peptid 4.8%, activ 4.4%, alpha 4.2%, kinas 3.7%, protein 2.7%, ca2 2.4%, phosphoryl 2.1%, inhibit 2.0%, psa 1.8%, glycodelin 1.7%, integrin 1.4%, beta 1.2%, alpha.beta 1.1%, jh2 0.9%] Focused on biomedical research on peptide binding.
- Cluster 125 (110) [receptor 37.5%, bind 4.2%, agonist 3.3%, antagonist 2.2%, ligand 2.0%, nmda 1.2%, rat 1.2%, releas 1.0%, histamin 1.0%, orexin 0.9%, glutam 0.9%, gaba 0.9%, activ 0.9%, brain 0.9%, express 0.8%] Focused on biochemistry, particularly of (antagonist) receptors and their biding ability.
- Cluster 153 (80) [rat 33.8%, ethanol 8.6%, tcdd 2.7%, dopamin 1.4%, alcohol 1.1%, anim 0.9%, induc 0.8%, ahr 0.7%, express 0.6%, exposur 0.6%, receptor 0.6%, liver 0.6%, gene 0.5%, wistar 0.5%, morphin 0.5%] Focused on genetic research, testing the influence of ethanol, dopamine, and alcohol, on biomedical models such as winstar rats and sprag dawley rats.
- Cluster 194 (70) [inhibit 5.2%, pdgf 3.3%, hif 2.9%, inhibitor 2.7%, liver 2.6%, activ 1.7%, enzym 1.4%, vitro 1.3%, allograft 1.3%, oxid 1.1%, tg2 1.0%, mum 1.0%, prodrug 1.0%, microsom 0.8%, rat 0.8%] Focused on inhibitors to platelet derived growth factor (pdgf).
- Cluster 86 (44) [neuron 26.8%, gdnf 6.9%, kcc2 2.3%, brain 1.7%, gabaerg 1.5%, axon 1.5%, gaba 1.3%, activ 1.3%, kinas 1.1%, ret

1.0%, merlin 1.0%, tau 1.0%, cell 0.9%, gfralpha1 0.9%, hippocamp 0.6%] Focused on the effect of particular proteins in the normal development of the brain., particularly Brain-Derived Neurotrophic Factor (BDNF) and neurons known to develop neurofibrillary tangles (NFTs).

- Cluster 141 (98) [mice 43.4%, transgen 3.0%, alpha 1.7%, cartilag 1.7%, express 1.0%, collagen 1.0%, knockout 1.0%, transgenic.mice 0.9%, adrenoceptor 0.8%, trkb 0.6%, receptor 0.5%, alpha.adrenoceptor 0.5%, knockout.mice 0.5%, type 0.5%, neuron 0.4%] Focused on various clinical trials using transgenic mice, collagen (from calcified and non-calcified cartilage).

2.1.1.2. Biomolecular Complexes

- Cluster 154 (67) [cell 12.3%, integrin 6.6%, adhes 4.7%, membran 4.4%, actin 2.6%, leukocyt 1.4%, collagen 1.4%, osteoclast 1.2%, chymas 1.2%, laminin 1.1%, beta 1.0%, phosphoryl 1.0%, golgi 0.9%, mediat 0.9%, smc 0.8%] Focused on the biological process cell adhesion and the extracellular matrix, with emphasis on integrins.
- Cluster 45 (54) [cell 15.7%, apoptosi 14.0%, caspas 9.7%, cell.death 9.3%, death 6.5%, apoptot 2.3%, induc 1.8%, activ 1.2%, mediat 0.7%, inhibitor 0.6%, protein 0.5%, nuclear 0.5%, neuron 0.4%, numa 0.4%, hsp60 0.4%] Focused on apoptosis / cell death, and caspase enzymes.
- Cluster 157 (159) [cell 59.3%, express 1.6%, line 1.5%, cell.lines 1.2%, cancer 0.8%, cell.line 0.6%, prolifer 0.5%, tumor 0.5%, vitro 0.5%, human 0.4%, protein 0.4%, gene 0.4%, receptor 0.3%, cultur 0.3%, neuron 0.3%] Focused on research utilizing cell lines, with emphasis on (DNA / RNA) expressions.
- Cluster 195 (83) [cell 26.3%, mast 3.1%, epitheli 2.4%, cultur 2.2%, mast.cells 1.8%, c1p 0.7%, respons 0.6%, clodron 0.6%, fluoresc 0.6%, protein 0.6%, bacteri 0.6%, growth 0.5%, caco 0.5%, calcium 0.5%, count 0.5%] Focused on mast cells, cell cultures, and epithelial cells.
- Cluster 4 (42) [vegf 39.2%, lymphat 7.3%, vessel 5.4%, endotheli 4.6%, vegfr 4.2%, vascular 3.2%, angiogenesi 1.6%, vascular.endothelial 1.1%, vegf.vegf 1.1%, growth 1.1%, cell 1.0%, endothelial.growth 1.0%, vascular.endothelial.growth 0.9%, express 0.9%, tumor 0.9%] Focuses on the effect of the vascular endothelial growth factor on the blood and lymph systems.

- Cluster 176 (118) [express 17.2%, tumor 7.6%, carcinoma 6.7%, tumour 3.1%, cell 2.9%, cancer 2.3%, stain 2.2%, cox 1.7%, tissu 1.7%, surviv 1.1%, tenascin 0.9%, patient 0.9%, epithelium 0.8%, keratinocyt 0.8%, hyaluronan 0.7%] Focused on tissue growth and (cancerous) cell reproduction, with emphasis on (DNA / RNA) expressions.
- Cluster 44 (46) [ifn 18.7%, ifn.gamma 8.8%, gamma 5.0%, express 5.0%, cell 4.0%, cytokin 3.5%, immun 1.9%, mrna 1.4%, lymphocyt 1.4%, stat1 1.2%, induc 1.1%, th1 1.1%, spore 1.0%, respons 0.9%, ccr5 0.9%] Focused on the body's immune response to various stimulants, with emphasis on (DNA / RNA) expressions.
- Cluster 174 (72) [express 9.6%, cell 5.5%, mrna 5.4%, tissu 4.3%, gata 2.3%, snurf 2.3%, protein 1.8%, insulin 1.4%, gdf 1.4%, mesenchym 1.3%, regul 1.1%, bmp 1.1%, igf 1.0%, inhibin 0.9%, adipos 0.9%] Focused on various methods for analyzing RNA and DNA, with emphasis on RNA/DNA expressions.
- Cluster 173 (120) [express 14.1%, cell 13.1%, gene 6.8%, transcript 4.7%, protein 4.7%, regul 3.3%, promot 1.8%, activ 1.5%, prostat 1.0%, gene.expression 1.0%, human 0.7%, mrna 0.7%, pim 0.6%, 1alpha 0.6%, induc 0.6%] Focused on proteins that regulate various cell activities.
- Cluster 144 (92) [gene 31.6%, chromosom 7.8%, express 4.5%, microarrai 1.7%, cancer 1.4%, amplif 1.3%, copy.number 1.1%, cgh 1.0%, protein 1.0%, copi 1.0%, overexpress 0.9%, tumor 0.9%, gene.expression 0.9%, transcript 0.8%, genom 0.8%] Focused on root causes of particular genetic disorders.

2.1.2. Genetic Science

2.1.2.1. Genetics of Cancerous Cells

- Cluster 74 (72) [linkag 10.1%, chromosom 6.6%, marker 5.8%, famili 5.5%, loci 5.4%, genet 4.3%, locu 4.2%, gene 3.2%, region 2.6%, microsatellit 2.0%, genom 1.9%, popul 1.8%, haplotyp 1.4%, qtl 1.2%, suscept 1.2%] Focused on genome wide scans / genome wide linkages to various health conditions, with emphasis on chromosomes, gene markers, and gene families..
- Cluster 143 (105) [polymorph 14.3%, genotyp 11.4%, gene 8.5%, allel 7.9%, snp 5.8%, haplotyp 2.0%, genet 1.8%, nucleotid 1.3%, popul 1.2%, dna 1.2%, exon 1.0%, single.nucleotide 0.9%, suscept 0.7%,

diabet 0.7%, frequenc 0.7%] Focused on genetic variations such as single nucleotide polymorphisms.

- Cluster 119 (138) [mutat 50.2%, gene 4.9%, famili 2.6%, patient 1.3%, cancer 1.2%, exon 1.0%, phenotyp 0.8%, tumor 0.7%, sequenc 0.7%, diseas 0.7%, protein 0.6%, mtdna 0.6%, cll 0.6%, colorect 0.6%, delet 0.5%] Focused genetic makeup of cancerous cells and their hereditary nature, with particular emphasis on gene mutations and gene families.

2.1.2.2. Cancer Risk Factors and Detection

- Cluster 2 (55) [mmp 73.0%, timp 1.5%, metalloproteinas 1.4%, express 1.3%, matrix 1.0%, mmp.mmp 1.0%, cell 0.7%, mt1.mmp 0.5%, mt1 0.5%, collagen 0.4%, tumor 0.4%, matrix.metalloproteinase 0.4%, tissu 0.3%, mmp.expression 0.3%] Focused on various matrix metalloproteinases (MMPs) enzymes and their implication on cancer growth.
- Cluster 69 (94) [cancer 36.0%, breast 5.4%, breast.cancer 4.7%, risk 4.5%, incid 2.0%, women 1.6%, cancer.risk 1.6%, prostat 1.6%, screen 1.3%, prostate.cancer 1.2%, case 1.1%, cohort 0.7%, lung.cancer 0.6%, control 0.6%, year 0.6%] Focused on cancer, particularly breast cancer, clinical trials identifying common risk factors.
- Cluster 98 (49) [cancer 21.3%, patient 10.2%, breast 4.8%, cancer.patients 3.6%, breast.cancer 3.4%, prostat 1.9%, serum 1.4%, prognost 1.1%, stage 0.9%, thyroid 0.9%, cea 0.9%, prostate.cancer 0.9%, treatment 0.8%, surviv 0.8%, breast.cancer.patients 0.8%] Focused on measurement techniques for (early) detection of cancer.

2.2. Medical Treatments

2.2.1. Treatment of Medical Conditions

2.2.1.1. Surgical Treatment, Medicines

- Cluster 66 (53) [surviv 13.9%, patient 12.5%, docetaxel 4.8%, chemotherapi 3.4%, median 3.2%, cancer 2.0%, melanoma 1.7%, month 1.7%, year 1.6%, uveal 1.0%, radiotherapi 0.9%, arm 0.9%, uveal.melanoma 0.8%, median.survival 0.8%, treatment 0.8%] Focused on cancer patients, particularly lung cancer survivors, and the treatments they took including chemotherapy treatments such as docetaxel.
- Cluster 50 (42) [transplant 18.1%, patient 12.9%, cmv 4.1%, liver 2.3%, mobilis 2.1%, remiss 1.8%, relaps 1.6%, surviv 1.6%, psc 1.5%, allograft 1.0%, cell 1.0%, regimen 0.9%, therapi 0.9%, donor 0.9%,

graft 0.8%] Focused patient survival of organ transplant surgery, particularly liver transplants.

- Cluster 58 (41) [pylori 21.1%, patient 5.5%, coeliac 4.3%, gastriti 4.1%, diseas 1.9%, coeliac.disease 1.8%, ulcer 1.8%, clozapin 1.3%, gastric 1.3%, erad 1.2%, tuberculosi 1.1%, duoden 1.1%, antibodi 0.9%, mucosa 0.9%] Focused on bacteria known to cause ulcers and other stomach related diseases.
- Cluster 184 (90) [patient 17.8%, infect 5.0%, pneumonia 3.2%, pancreat 2.5%, crp 2.5%, antibodi 2.2%, serum 2.2%, urinari 1.6%, acut 1.4%, acute.pancreatitis 1.0%, diagnosi 1.0%, level 0.8%, tract 0.8%, sever 0.8%, diagnost 0.8%] Focused on treatments for infectious diseases, particularly those that effect the lungs, the urinary tract, or the pancreas.
- Cluster 190 (155) [patient 48.1%, year 1.7%, diseas 1.4%, arthriti 1.4%, diabet 1.2%, control 0.8%, clinic 0.8%, joint 0.6%, group 0.5%, epilepsi 0.5%, seizur 0.5%, rheumatoid 0.5%, ag 0.4%, renal 0.4%, diabetic.patients 0.3%] Focused on genetic analysis of diseases such as arthritis, epilepsy, and diabetes.
- Cluster 179 (68) [patient 15.9%, mri 4.2%, atrophi 3.3%, syndrom 3.0%, diseas 1.9%, brain 1.8%, alzheim 1.3%, ftd 1.2%, hear 1.1%, abnorm 1.0%, imag 1.0%, clinic 0.9%, alzheimer.disease 0.9%, lesion 0.8%, exfoli 0.7%] Focused on diagnosis and etiology of alzheimer's disease.
- Cluster 16 (30) [node 23.2%, sentinel 10.7%, tumour 4.4%, sentinel.node 3.3%, patient 3.0%, metastas 1.9%, lymph 1.9%, sentinel.nodes 1.9%, axillari 1.8%, sln 1.5%, tumor 1.4%, metastasi 1.3%, axilla 1.0%, breast 1.0%, breast.cancer 1.0%, lymph.node 0.9%, snb 0.9%, cancer 0.7%, malign 0.7%, lymphoscintigraphi 0.7%, biopsi 0.7%] Focused on the use of sentinel node biopsy in the treatment of cancers.
- Cluster 43 (42) [pain 46.1%, patient 5.9%, surgeri 1.1%, group 1.1%, treatment 0.9%, tonsillectomi 0.8%, disabl 0.7%, analgesia 0.7%, chronic 0.6%, sensori 0.5%, back.pain 0.5%, symptom 0.5%, ketoprofen 0.5%, month 0.5%, mouth 0.4%, back 0.4%] Focused on pain intensity (e.g., postoperative, injury, etc.) and pain relief.
- Cluster 106 (51) [fractur 15.9%, knee 9.7%, patient 7.0%, hip 3.4%, joint 1.7%, arthroplasti 1.4%, disloc 1.4%, implant 1.3%, year 1.1%, complic 0.9%, screw 0.8%, month 0.7%, mri 0.6%, pain 0.6%, leg 0.6%] Focused on treatments for orthopedic injuries.

- Cluster 181 (87) [patient 13.7%, surgeri 7.5%, postop 4.4%, oper 2.7%, group 2.3%, complic 2.2%, incontin 2.2%, hospit 1.7%, month 1.5%, injuri 1.2%, surgic 1.1%, year 1.1%, anaesthesia 0.9%, amput 0.9%, ey 0.9%] Focused on postoperative complications.

2.2.1.2. Medicines and Medical Treatment Equipment

- Cluster 110 (58) [patient 9.9%, icu 5.9%, aneurysm 5.2%, postop 4.2%, hospit 3.5%, qol 2.8%, score 2.6%, cabg 2.5%, mortal 2.0%, sah 1.6%, care 1.3%, intensive.care 1.3%, admiss 1.1%, fistula 1.0%, cost 0.9%] Focused on hospital intensive care unit procedures and postoperative recovery, particularly for those recovering from aneurysm surgery.
- Cluster 36 (44) [coronari 12.3%, arteri 11.8%, coronary.artery 5.5%, graft 4.8%, bypass 3.3%, reperfus 3.2%, myocardi 2.4%, cabg 1.9%, patient 1.7%, artery.bypass 1.7%, pump 1.6%, group 1.5%, coronary.artery.bypass 1.4%, postop 1.3%, precondit 1.2%] Focused on coronary artery bypass technologies and procedures.
- Cluster 41 (31) [galantamin 7.5%, vad 5.7%, cognit 5.2%, month 4.8%, patient 3.7%, therapi 2.1%, alzheim 1.9%, tacrolimu 1.9%, dementia 1.9%, bpa 1.7%, alzheimer.disease 1.7%, rivastigmin 1.6%, donepezil 1.6%, cadi 1.6%, bnct 1.3%] Focused on treatments for cerebrovascular diseases, such as alzheimer’s disease and dementia, with galantanim.
- Cluster 32 (36) [stroke 32.7%, seizur 9.3%, patient 4.4%, depress 2.7%, ect 2.4%, ischem 1.7%, ischemic.stroke 1.4%, month 1.1%, cognit 0.9%, stroke.patients 0.9%, infarct 0.8%, trial 0.8%, sever 0.7%, scale 0.6%, score 0.6%] Focused on seizures and ischemic strokes.
- Cluster 133 (53) [patient 9.6%, cardiac 8.0%, death 2.8%, ami 2.3%, guidelin 2.2%, hospit 2.0%, drug 1.8%, cardiac.arrest 1.8%, risk 1.7%, resuscit 1.6%, infarct 1.5%, defibril 1.3%, myocardial.infarction 1.2%, arrest 1.2%, myocardi 1.2%] Focused on cardiac arrest resuscitation equipment such as defibrillation electrodes, and clinical drugs.
- Cluster 113 (46) [orbit 11.5%, rotat 5.5%, asteroid 5.2%, lunar 4.9%, rai 2.2%, solar 2.0%, period 2.0%, lightcurv 1.5%, earth 1.3%, mission 1.2%, instrument 1.0%, cix 1.0%, accret 0.9%, itokawa 0.8%, neptun 0.8%] Focused on astronomy, especially asteroid orbits and lunar orbits.

- Cluster 67 (52) [pressur 17.8%, blood.pressure 11.9%, blood 6.0%, mmhg 4.4%, hypertens 3.5%, systol 2.7%, levosimendan 2.6%, heart 1.9%, heart.rate 1.5%, arteri 1.1%, diastol 1.0%, patient 0.9%, baroreflex 0.9%, insulin 0.7%, hrv 0.6%] Focused on the cardiovascular system, blood pressure and treatments for these conditions.
- Cluster 111 (56) [blood.flow 8.7%, flow 7.8%, arteri 6.0%, blood 5.7%, perfus 5.4%, carotid 1.6%, patient 1.2%, stenosi 1.2%, myocardi 1.1%, cerebr 1.1%, diabet 1.1%, coronari 0.8%, veloc 0.7%, cbv 0.7%, min 0.7%] Focused on cardiovascular / circulatory system and blood flow.
- Cluster 104 (42) [plasma 5.0%, concentr 3.1%, gemfibrozil 2.7%, celiprolol 2.6%, pravastatin 2.4%, hour 2.3%, auc 2.2%, repaglinid 2.2%, mug 2.1%, simvastatin 1.9%, pharmacokinet 1.6%, rifampicin 1.6%, assai 1.5%, dose 1.4%, ropivacain 1.3%] Focused on clinical trials of various prescription cholesterol drugs, especially gemfibrozil, celiprolol, and pravastatin.
- Cluster 94 (42) [dose 31.7%, iop 6.7%, mifepriston 2.0%, radiat 1.5%, treatment 1.4%, ey 1.2%, tonomet 1.2%, mmhg 1.2%, mbq 1.1%, group 1.0%, ospemifen 1.0%, dap 0.9%, mgy 0.8%, vagin 0.8%, misoprostol 0.8%] Focused on clinical trials using mifepristone for abortions.
- Cluster 162 (85) [treatment 12.8%, patient 10.7%, placebo 9.5%, dose 2.6%, group 2.3%, symptom 1.6%, trial 1.3%, advers 1.1%, toler 1.1%, week 0.9%, asthma 0.8%, treat 0.8%, therapi 0.7%, month 0.7%, double.blind 0.7%] Focused on clinical trials using drugs and placebos (for various health conditions such as asthma, etc.) to ascertain effectiveness.
- Cluster 197 (86) [group 13.6%, placebo 5.4%, insulin 4.8%, glucos 2.6%, treatment 2.0%, min 1.9%, propofol 1.8%, blood 0.9%, patient 0.8%, diabet 0.8%, sevofluran 0.6%, subject 0.6%, control 0.6%, uptak 0.5%, concentr 0.5%] Focused on clinical trials using medicines (primarily insulin), and placebos for various conditions, but particularly for stabilizing blood glucose levels.

2.2.2. Lifestyle, Pregnancy, and Neuropsychological

2.2.2.1. Healthcare, Long Term Medical Conditions, Mental Disorders

- Cluster 52 (51) [muscl 28.4%, exercis 23.0%, emg 3.3%, train 1.0%, isometr 0.8%, perfus 0.7%, paraspin 0.7%, forc 0.7%, patient 0.6%, glucose.uptake 0.6%, ssc 0.6%, exercise.induced 0.4%, knee 0.4%, ssc.exercise 0.4%, fatigu 0.4%] Focuses on various aspects of physical activity / exercise.
- Cluster 7 (33) [train 40.4%, strength 6.6%, group 2.9%, muscl 2.6%, exercis 2.2%, strength.training 1.9%, maxim 1.9%, neck 1.1%, training.period 0.9%, resistance.training 0.8%, endur 0.8%, training.group 0.8%, control 0.6%, patient 0.6%, exercise.training 0.6%, control.group 0.5%, isometr 0.5%] Focuses on various aspects of physical training including physical therapy.
- Cluster 57 (70) [bone 44.1%, bmd 3.5%, implant 2.1%, fractur 1.1%, group 1.0%, alendron 0.9%, femor 0.9%, bone.mineral 0.8%, graft 0.8%, screw 0.8%, fixat 0.7%, miner 0.7%, month 0.7%, glenoid 0.5%, femoral.neck 0.5%] Focused on the skeletal system and various fracture healing techniques.
- Cluster 73 (48) [women 22.2%, hrt 13.9%, menopaus 3.2%, bone 3.0%, postmenopaus 1.4%, ag 1.3%, hysterectomi 1.2%, year 1.1%, bmd 0.9%, treatment 0.9%, iu 0.9%, lng 0.9%, cervic 0.8%, postmenopausal.women 0.8%, lng.ius 0.7%] Focused on clinical studies of menopausal women.
- Cluster 132 (55) [intak 14.5%, vitamin 5.8%, diet 4.6%, dietari 3.5%, folat 2.3%, lignan 2.1%, fat 1.7%, food 1.4%, plasma 1.3%, isoflavon 1.2%, concentr 1.1%, urin 1.1%, acid 1.1%, subject 1.1%, group 1.0%] Focuses on dietary patterns and nutritional contents of foods.
- Cluster 121 (44) [fish 5.2%, plasma 4.8%, concentr 4.6%, leptin 4.5%, feed 4.4%, cortisol 2.9%, fat 2.5%, fed 2.3%, diet 2.2%, ghrelin 1.2%, contaminated.grains 1.2%, testosteron 1.2%, fast 1.1%, bodi 1.0%, group 1.0%] Focused on biological response of fish to drastic dietary change, particularly feed reduction and fasting. Also focused on proteins such as leptin, and hormones that regulate weight and fat concentration for humans.
- Cluster 53 (72) [cholesterol 23.6%, ldl 6.4%, hdl 4.5%, lipoprotein 4.2%, sterol 3.3%, serum 3.2%, lipid 1.9%, density.lipoprotein 1.9%, triglycerid 1.0%, plasma 0.9%, hdl.cholesterol 0.9%, concentr 0.9%, ldl.cholesterol 0.8%, plant 0.8%, low.density.lipoprotein 0.8%] Focused on the effect of clinical drugs, diet / lifestyle on ldl and hdl cholesterol levels.

- Cluster 82 (66) [insulin 10.6%, obes 8.2%, weight 5.8%, weight.loss 3.5%, metabolic.syndrome 3.1%, metabol 2.9%, bmi 2.2%, glucos 2.1%, pco 2.1%, syndrom 1.9%, women 1.4%, loss 1.4%, men 1.2%, subject 1.1%, bodi 1.1%] Focused on the correlation between blood sugar / glucose levels and obesity.
- Cluster 64 (92) [diabet 39.5%, type.diabetes 8.6%, insulin 3.4%, glucos 3.0%, type 2.6%, risk 2.1%, subject 1.4%, hla 1.1%, children 0.6%, genotyp 0.6%, glucose.tolerance 0.6%, ldl 0.5%, autoantibodi 0.5%, women 0.4%, dqb1 0.4%] Focused on diabetes, with particular emphasis in production of insulin and its effect on blood glucose levels.
- Cluster 11 (36) [chd 32.3%, coronari 6.2%, risk 3.0%, coronary.heart 2.9%, coronary.heart.disease 2.8%, heart.disease 2.6%, diseas 1.9%, heart 1.9%, men 1.9%, mortal 1.7%, ag 1.3%, risk.factors 1.3%, women 1.1%, disease.chd 1.0%, death 1.0%, heart.disease.chd 1.0%] Focused causes and risk factors for heart disease.
- Cluster 80 (61) [mortal 26.4%, death 9.2%, stroke 6.0%, men 4.2%, ag 2.9%, risk 2.7%, women 2.5%, cardiovascular 1.0%, socioeconom 0.9%, diseas 0.9%, year 0.9%, caus 0.8%, alcohol 0.7%, popul 0.6%, declin 0.6%] Focuses on mortality risk factors, particularly for risk of stroke or cardiovascular diseases.
- Cluster 187 (119) [men 19.6%, women 7.3%, ag 5.7%, alcohol 3.8%, smoke 2.5%, year 2.5%, men.women 2.1%, subject 1.2%, popul 1.0%, cdt 1.0%, disabl 0.9%, consumpt 0.8%, bodi 0.8%, preval 0.6%, obes 0.6%] Focused on comparing the effects of lifestyles (smoking, alcohol consumption) on the health of men and women.
- Cluster 196 (74) [ag 6.1%, social 3.3%, peopl 2.9%, person 2.5%, self 1.8%, group 1.5%, unemploy 1.5%, year 1.4%, men 1.4%, food 1.3%, life 1.2%, driver 1.2%, countri 1.1%, perceiv 1.1%, posit 0.9%] Focused on socioeconomic factors in Europe and Finland.
- Cluster 151 (45) [balanc 5.4%, elderli 5.1%, postur 4.2%, subject 3.8%, train 2.9%, dentur 2.5%, mood 2.3%, intervent 1.7%, dentist 1.6%, ag 1.3%, test 1.2%, psychotrop 1.1%, denturist 1.0%, year 0.9%, functional.balance 0.9%] Focused on physical and mental health issues, particularly balance and posture issues for the elderly.
- Cluster 145 (69) [exposur 21.7%, risk 15.2%, worker 4.4%, cancer 2.1%, mortal 2.0%, asthma 1.9%, occup 1.7%, asbesto 1.3%, smoke 1.0%, case 0.8%, lung 0.7%, bitumen 0.7%, lung.cancer 0.6%, level 0.5%, asphalt 0.5%] Focused on risk factors for lung cancer and asthma, such as smoking, exposure to asbestos, etc.

- Cluster 137 (52) [symptom 23.9%, asthma 5.0%, pain 3.1%, tmd 2.1%, exposur 1.7%, subject 1.4%, questionnair 1.4%, health 1.3%, injuri 1.3%, athlet 0.9%, occup 0.7%, sport 0.7%, depress 0.7%, year 0.7%, physic 0.6%] Focused on classifying / identifying symptoms for medical conditions, particularly temporomandibular disorder (TMD), and asthma.
- Cluster 108 (52) [work 29.5%, employe 5.0%, sick 2.2%, organiz 1.7%, health 1.6%, stress 1.4%, team 1.1%, voic 1.1%, vocal 1.1%, symptom 1.0%, worker 0.8%, occup 0.7%, justic 0.7%, mental 0.7%, questionnair 0.7%] Focused on organization cultures, climates, and employee health.
- Cluster 30 (40) [adolesc 36.8%, smoke 5.4%, school 5.1%, health 3.6%, tobacco 1.9%, behaviour 1.4%, parent 1.2%, girl 1.1%, boi 1.0%, alcohol 1.0%, self 0.9%, survei 0.9%, drink 0.8%, ag 0.7%, adulthood 0.7%,] Focused on (lifestyle) factors and behaviors contributing to the health of Finnish adolescents, especially smoking (tobacco) and alcohol.
- Cluster 112 (90) [health 42.9%, care 3.6%, health.care 2.4%, servic 2.2%, incom 1.0%, survei 0.9%, self 0.9%, alcohol 0.8%, physician 0.7%, rated.health 0.7%, self.rated.health 0.7%, intervent 0.6%, self.rated 0.6%, dental 0.6%, municip 0.5%] Focused on the health care system in Finland.
- Cluster 26 (40) [nurs 53.7%, care 5.5%, ethic 2.5%, envi 1.5%, health 1.4%, empower 1.3%, physician 0.9%, work 0.7%, compet 0.6%, learn 0.6%, sow 0.6%, medic 0.5%, practic 0.4%, intercultur 0.4%, person 0.4%] Focused on healthcare and health welfare in Finland, with emphasis on nursing and ethics..
- Cluster 79 (49) [patient 18.1%, nurs 17.5%, care 9.3%, ethic 2.4%, hospit 1.7%, dementia 1.1%, physician 1.0%, percept 0.9%, nursing.care 0.9%, clinic 0.9%, am 0.8%, record 0.8%, cari 0.8%, dnar 0.7%, famili 0.7%] Focuses on dementia, and available nursing and patient care in Finland.
- Cluster 1 (23) [suicid 69.8%, attempt 2.1%, suicide.attempters 2.0%, psychiatr 1.6%, disord 1.1%, depress 1.0%, drown 0.7%, psychiatric.consultation 0.6%, consult 0.6%, mood 0.6%] Focuses on assessing and identifying factors (e.g., medical disorders, mental disorders, etc.) that can be correlated with risk for suicide.
- Cluster 90 (65) [depress 14.4%, disord 11.8%, schizophrenia 5.5%, offend 2.8%, violent 2.8%, hopeless 2.6%, crimin 2.4%, psychiatr 1.9%, sleep 1.2%, risk 1.1%, mental 1.1%, symptom 1.1%, adolesc

1.0%, general.population 0.8%, subject 0.8%] Focused on depression, schizophrenia and other mental disorder.

- Cluster 25 (28) [schizophrenia 31.9%, patient 7.2%, dsm 4.4%, neurolept 3.6%, delirium 1.6%, schizophrenia.patients 1.5%, schizophren 1.2%, psychot 1.0%, dsm.iii 0.9%, notch4 0.9%, criteria 0.9%, dialogu 0.9%, patients.schizophrenia 0.9%, schizophrenic.patients 0.8%, csf 0.6%] Focused on schizophrenia and other central nervous system disorders.

2.2.2.2. Pregnancy and Neuropsychological

- Cluster 105 (59) [pregnanc 25.1%, women 6.5%, twin 6.4%, abort 3.2%, embryo 3.2%, rate 2.4%, birth 1.9%, deliveri 1.8%, gestat 1.0%, outcom 0.9%, group 0.7%, pregnant 0.7%, rd 0.6%, preeclampsia 0.6%, regist 0.6%] Focuses on various aspects of pregnancy and child birth in Finland.
- Cluster 89 (37) [infant 18.8%, gestat 3.6%, gestational.age 1.9%, birth 1.8%, cortisol 1.8%, week 1.4%, cord 1.4%, preterm 1.3%, weight 1.3%, bpd 1.2%, serum 1.2%, ag 1.1%, deliveri 1.1%, igfbp 1.1%, pregnanc 1.1%] Focused on human pregnancies and various factors effecting fetal development.
- Cluster 8 (30) [vaccin 37.9%, fluorid 7.6%, serotyp 6.5%, pneumococc 4.3%, immun 2.5%, conjug 1.8%, antibodi 1.2%, infant 0.9%, fluorosi 0.9%, aom 0.8%, 19f 0.8%, valent 0.7%, opa 0.7%, toothpast 0.7%, polysaccharid 0.7%, pneumococcal.conjugate 0.5%, gmc 0.5%, pnc 0.5%, bcg 0.5%, hepat 0.4%] Focuses on vaccines for pneumococcal serotype.
- Cluster 21 (33) [aom 11.3%, otiti 9.3%, otitis.media 8.9%, media 5.0%, pneumonia 4.7%, children 3.5%, acute.otitis 3.3%, acute.otitis.media 3.1%, acut 2.7%, ag 2.1%] Focused on the treatment of an ear infection disease that is prevalent to children.
- Cluster 63 (76) [parent 21.0%, child 13.1%, mother 10.5%, children 8.4%, famili 8.1%, infant 2.0%, father 1.8%, matern 1.1%, behavior 0.7%, pain 0.6%, social 0.5%, girl 0.4%, percept 0.4%, depress 0.4%, interact 0.4%] Focuses on sociological aspects of family life.
- Cluster 135 (133) [children 56.3%, ag 2.9%, year 1.6%, school 1.5%, group 0.8%, test 0.5%, child 0.5%, infect 0.5%, asthma 0.4%, control 0.4%, wheez 0.3%, born 0.3%, respiratori 0.3%, girl 0.3%, enteroviru 0.3%] Focused on health, medicines, etc, for school aged children.

- Cluster 49 (39) [children 12.2%, read 7.0%, erp 4.9%, auditori 4.0%, mmn 3.7%, sound 3.3%, infant 3.0%, ag 2.1%, tone 1.6%, school 1.4%, teeth 1.3%, year 1.2%, cari 1.2%, sensori 0.9%] Focused on the development of children's neural systems and their language functions.
- Cluster 46 (51) [mmn 13.7%, sound 9.7%, auditori 6.9%, stimulu 2.9%, stimuli 2.6%, deviant 2.1%, erp 2.0%, speech 1.7%, visual 1.6%, process 1.3%, attent 1.2%, face 1.2%, mismatch 1.2%, tone 1.1%, cortic 1.1%, percept 1.0%] Focused on the neural system's underlying language functions.
- Cluster 142 (83) [auditori 8.8%, left 4.2%, cortex 4.0%, stimuli 2.9%, tone 2.6%, respons 2.4%, right 2.0%, cortic 2.0%, ear 2.0%, hemispher 1.8%, subject 1.8%, stimulu 1.7%, brain 1.5%, activ 1.5%, sound 1.2%] Focused on the cortex, and methods for visualizing brain activity such as magnetic resonance imaging.
- Cluster 42 (51) [word 46.0%, languag 7.5%, queri 4.3%, letter 1.0%, syllabl 1.0%, name 1.0%, translat 0.9%, string 0.8%, read 0.8%, dictionari 0.7%, english 0.7%, infinite.words 0.6%, infinit 0.6%, lexic 0.6%, sentenc 0.6%] Focused on linguistics / phonological and semantic processes of words.
- Cluster 20 (40) [host 36.7%, parasit 18.7%, speci 6.0%, larva 1.3%, fish 1.2%, grb 0.7%, plant 0.7%, parasitoid 0.7%, host.plant 0.6%, paranoplocephala 0.6%, butterfli 0.6%, beetl 0.5%, genu 0.5%, parasite.species 0.5%, popul 0.5%] Focused on ecological species.
- Cluster 0 (13) [ternari 12.4%, ternary.system 8.1%, thermodynam 4.8%, system.optimized.experimental 2.2%, thermodynamic.ternary.system 2.2%, sgte 2.2%, system.optimized 2.2%, ternary.system.optimized 2.2%, optimized.experimental 2.2%, thermodynamic.ternary 2.2%, phase.equilibrium 2.2%, thermodynamic.parameters 2.1%, equilibrium.data.ternary 1.8%, phase.equilibrium.data 1.8%] Focuses on equilibrium thermodynamic descriptions of (copper) ternary material systems.
- Cluster 17 (32) [sleep 51.9%, spindl 3.7%, nerv 1.5%, apnea 1.3%, nasal 0.9%, wake 0.9%, rem 0.7%, wave 0.7%, subject 0.7%, patient 0.6%, eeg 0.6%, apnoea 0.5%, night 0.5%, nrem 0.4%, rem.sleep 0.4%] Focused on sleep behavior and sleep disorders, and measurement of sleep spindle.